A data-driven approach to estimate battery cell temperature using a nonlinear autoregressive exogenous neural network model

Md Mehedi Hasan[a], Seyyed Ali Pourmousavi Kani[b,*], Ali Jahanbani Ardakani[c], Tapan Kumar Saha[a]

[a] *School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane, QLD 4072, Australia*
[b] *The School of Electrical and Electronic Engineering, The University of Adelaide, Adelaide, SA 5005, Australia*
[c] *Department of Electrical and Electronic Engineering, Iowa State University, IA 50011, United States*

Abstract

Battery cell temperature is a key parameter in battery life degradation, safety, and dynamic performance. Intense charging-discharging operations and high-ambient temperatures escalate battery cell temperature, which in turn accelerates its degradation. Therefore, accurate battery cell temperature estimation can play a significant role in ensuring the optimal operation of a battery energy storage system (BESS). In order to estimate battery cell temperature as accurate as possible, use of non-linear models is imperative due to the non-linear nature of the battery operation. This paper proposes a data-driven model based on a Non-linear Autoregressive Exogenous (NARX) neural network to estimate battery cell temperatures in a utility-scale BESS, considering strongly-correlated independent variables, e.g., charging-discharging current and ambient temperature. Due to different temperature and weather characteristics in each season, seasonal NARX models have also been derived and compared with the universal one. The proposed models' performance has been verified using the field data collected from a grid-connected BESS within a PV plant. The simulation results show high accuracy of the proposed model compared to the measured data for both seasonal and universal models without considering the complexity of the large-scale battery and container thermal dynamics. In particular, in more than 95% of the time, the estimated values yield root mean squared errors (RMSE) below 1 °C in different conditions, which confirms the validity and accuracy of the proposed model. Moreover, seasonal models show better performance with 18% to 50% less RMSE on average (for 1 hour to 24 hours forward estimation) compared to the universal model.

*Corresponding author
Email addresses: mdmehedi.hasan@uq.net.au (Md Mehedi Hasan), a.pourm@adelaide.edu.au (Seyyed Ali Pourmousavi Kani), alij@iastate.edu (Ali Jahanbani Ardakani), saha@itee.uq.edu.au (Tapan Kumar Saha)

Keywords: Battery cell temperature, Battery Energy Storage System (BESS), Time Series NARX feedback neural network, PV plant, Seasonality

1. Introduction

With the increasing amount of renewable energy sources (RES) in the power system, energy storage devices (e.g., batteries) are becoming an integral part of the power system for safe and secure operation. They can support solar and wind energy production by storing the excess energy when available. Additionally, due to high variability and unpredictability in the RES generation profiles, the BESS can be applied to smooth the output power [1]. While the application of BESS in the modern power system is inevitable and storage technologies are improving quickly on economic and technical aspects, battery degradation remains a big concern [2, 3]. The capacity of batteries deteriorates due to charging and discharging activities as well as during idle conditions. Among various factors affecting battery ageing process, battery cell temperature is known to have a major effect by escalating battery degradation exponentially [4, 5] and intensifying safety concerns [6]. In fact, every 10 °C rise in the battery cell temperature doubles the battery degradation rate [7]. Surging battery cell temperature occurs because of charging-discharging operations and ambient temperature [8, 9]. Moreover, the accumulated heat in a utility-scale BESS that is placed in a confined container, elevates the temperature inside the battery container, which increases cooling cost by requiring air conditioning units to operate at a higher rate to maintain the temperature withing the desired range [6]. Therefore, it is important to estimate the battery cell temperature evolution in time with respect to a planned operation profile and ambient interactions to have a better evaluation of the incurred cost. In particular, accurate temperature estimation is important for battery energy management systems (EMS), which operate battery in an optimal way considering its degradation and cooling costs. Using a battery cell temperature estimation algorithm, an effective energy management strategy can be established with an accurate cell temperature model. It will help to prevent unnecessary battery operation by taking different factors into account. Therefore, the BESS lifetime can be extended, which positively affects economic benefits of the entire system.

A number of research has been conducted on the battery thermal management of electric vehicles (EVs) and hybrid electric vehicles (HEVs), where they proposed general models of

battery thermal behaviour in [10–12] or specialised models for cold weather or lower temperatures [6, 9]. However, there is no research to investigate thermal behaviour of a utility-scale BESS in connection with a PV plant and the main grid. In this paper, thermal behaviour of a battery with respect to charge-discharge activities and ambient weather conditions is studied. It is noticeable that in most of the studies, thermo-electric battery physical model and thermal-conductivity based approaches have been followed to analyse thermal behaviour of lithium-ion batteries [6, 13]. These approaches require laboratory-based experiments for a single battery module, considering various factors such as battery ambient temperature, age, state of charge (SOC) and other operational conditions. In [14], a complicated thermo-dynamical model is developed for sealed battery packs in EVs and HEVs in contact with a cooling agent. While the models are accurate, the application of such models could become overwhelmingly complicated in the utility-scale BESS due to the large number of battery modules. Moreover, some of these models typically need several internal parameters to be measured continuously, which might render the solution expensive and unscalable for the utility-scale applications. Detailed thermal modelling of the entire utility-scale BESS is time consuming as it requires complex thermal modelling and intensive computational requirements that are not feasible for the energy management applications. In other words, modelling the complex nature of thermal interactions of thousands of battery cells with each other and the larger surrounding area, and active/passive cooling mechanisms through thermodynamic models will be a computationally intractable approach.

Typically, detailed thermal models of a battery requires a number of sensors in different parts of the cell to measure the necessary parameters [15]. These models usually consider the complex nature of the electrochemical components and their interactions in a chemical environment to estimate the battery cell temperature evolution under different operation regimes [16]. Particularly, physical principles of the particles are considered in the literature to model thermal dynamics of a battery cell, which can be classified into detailed and simplified models, such as reduced equivalent electric circuit [17, 18] and reduced electro-thermal coupled [19]. They require precise measurements of the internal and external parameters in a battery cell to estimate its thermal behaviour, which is not practical in a utility-scale BESS. Calculating heat generation, conductive and convective heat transfers of a battery cell requires a good number of parameters' measurements and estimations in a battery container, which is impractical or

expensive for a large-scale BESS. In addition, accumulated measurement errors and unknown disturbances lead to further inaccuracy in the cell temperature model. Therefore, an easy-to-use practical battery thermal model is needed for utility-scale applications, which can deal with non-linearity of the underlying system. It must also be computationally inexpensive to be integrated within on-line applications, such as an EMS. Therefore, a physics-free data-driven modelling is preferred in this paper.

Artificial intelligence methods are widely used in various research fields to model the unknown or complicated physical systems based on measured input and output parameters, such as fuzzy systems, nonparametric regression, wavelets, neural network, or a combination of them which is known as hybrid systems [20–23]. These methods are extensively used to model non-linear black-box systems in different applications [24–26]. The studies conducted with these methods demonstrate that any complex nature of dynamics and non-linearity can be modelled effectively. Therefore, these methods are considered for modelling battery thermal dynamics in this study using available input and output data of the battery operation. In this procedure, historical data of the input and output parameters are used to train a model considering the most influential independent variables.

Another advantage of data-driven models over analytical methods is that they can deal with the time-varying internal parameters of the BESS by receiving new data and re-training. For instance, higher internal resistance of the battery cell is inevitable when the cell ages, which contributes to higher cell temperature by excessive heat generation during charging-discharging activities. For analytical methods to work accurately, internal resistance should be measured periodically by interrupting battery operation and spending money on measurement tools and personnel. In data-driven methods, however, new measurement of external parameters and re-training the model can modify it according to the new changes in the battery cell operation. It can be done continuously without requiring expensive procedures. Therefore, the overall accuracy of the model will improve substantially with this low-cost solution over the lifetime of the BESS.

In this study, a data-driven approach, namely a Non-linear Autoregressive Network with Exogenous inputs (NARX), is used to estimate lithium-polymer battery cell temperature. This study is a part of the research that has been reported in [27]. It is shown that NARX is

more effective in terms of learning the long temporal dependencies than the other complicated methods such as wavelets, nonparametric regression, and fuzzy models [28]. Instead of developing a model for a single cell in a controlled experimental setup, the proposed model created from the operation of thousands of cells affecting each other in a large-scale BESS. The main goal is to create a thermal model of the battery cells within a container equipped with passive and active cooling mechanisms, which can be utilised in the EMS. Also, the non-linear nature of the phenomenon (i.e., the cell temperature variations due to charge-discharge regime and thermal interactions among cells with outside and cooling mechanisms) can be effectively captured by the NARX model. Moreover, we tested seasonality hypothesis and we showed that a seasonal model will work more accurately. To the best of our knowledge, the application of NARX for day-ahead large-scale battery temperature estimation has not been reported in the literature considering active and passive cooling mechanisms. Despite ordinary neural networks, NARX can model temporal information in a given time series, which is quite important. In other words, not only the input features contain useful information for estimation, but also the changes in input/output over time can provide additional information to increase the accuracy of estimation, which can be modelled by NARX. An ordinary NN treats every sample individually; thus, it fails to model the temporal information.

As a black-box modelling approach, NARX does not need a physical model of the underlying system and the thermal interactions among different mediums and effective parameters. The model can be created by battery operation and ambient temperature historical values measured in a utility-scale PV-BESS plant. In this study, the best structure of the NARX model (i.e., number of neurons and feedback delay) is selected by sensitivity analysis. Minute-by-minute field data, collected from The University of Queensland Gatton campus PV-BESS plant, is used by dividing data into training and test sets for the NARX model development and evaluation, respectively. In addition, in an attempt to improve accuracy of the model and to evaluate seasonal impact on the cell temperature modelling, seasonal NARX models are developed using the training dataset and evaluation of each model is done using different set of seasonal data (i.e., test dataset) for each season. Finally, the results obtained from universal model are compared with the ones obtained from seasonal ones, where the latter outperformed the former.

In this study, autoregressive exogenous neural network approach is adopted to forecast

hourly battery cell temperature for up to 24 hours ahead window. Developed non-liner model is established and tested with field data. The simulation outcomes presented in this paper confirms the robustness of the estimation model. In addition, a comparison study is also presented in this paper to compare the performance of two different models for battery cell temperature estimation.

The rest of the paper is organised as follows: Section 2 explains the proposed methodology. Section 3 presents the system under study, and Section 4 explains a wide range of results and discussions. Finally, the paper is concluded in Section 5 and future works are outlined.

2. Data-Driven Battery Cell Temperature Model

The NARX approach based on ANN is employed for battery cell temperature modelling due to its capability for non-linearity modelling considering exogenous parameters and feedback signal. The NARX model is based on the non-linear autoregressive model, which is commonly used in time-series modelling. The model can accept exogenous parameters as well as a feedback from the output with a certain delay. In this study, exogenous parameters are the battery current and ambient temperature (weather temperature or outside of the container temperature). Battery current (i.e., charge-discharge and idle regime) has direct influence on the battery cell temperature because of the ohmic and non-ohmic internal losses [11]. Also, the ambient temperature can change the temperature inside the container, which essentially changes cell temperature with some delay. Small changes in the ambient temperature might lead to significant heat accumulation in the battery container and consequently in the battery cell during charging-discharging as well as idle conditions. It also contributes to lower efficiency of the battery plant by requiring cooling systems to operate intensively.

Battery cell temperature, similar to many other natural phenomena, does not change like a step function. As it was shown in [5], battery cell temperature starts increasing gradually after a charge or discharge incident due to heat convection delay, physical layout of cells in the container, and cooling system operation. Therefore, limited number of samples from the past can be helpful to estimate future output with higher accuracy, which requires a feedback from output in the model.

2.1. NARX Models

The NARX model can be mathematically represented as:

$$\begin{aligned} y(t) = & f[y(t-1), y(t-2), .., y(t-n_y), \\ & u(t-1), u(t-2), .., y(t-n_u), \\ & e(t-1), e(t-2), e(t-n_e)] + e(t) \end{aligned} \tag{1}$$

where, $y(k)$ is the output or dependent variable; $u(k)$ refers to the input parameters; $e(k)$ is the noise sequence; t denotes the discrete time-step; and n_u and n_y represents the number of time-step delays for input and output parameters, respectively. Non-linear mapping function, $f(\cdot)$, is generally unknown. This function can be approximated using different model structures, such as fuzzy logic-based models, neural networks, etc. In this work, it is represented by a standard multi-layer perceptron (MLP) network in this study. MLP is a dominant structure of learning any type of continuous non-linear mapping when bundled with proper training algorithms.

2.2. Artificial Neural Networks

Artificial neural networks are simple imitations of how human brain processes the data. They can be used for several applications such as classification, regression, and pattern recognition. Fig. 1 shows a typical neural network consisting of three layers, namely, input, hidden, and output layer.

An ANN model can be shown mathematically as follows:

$$\tilde{Y} = \Psi\left(\sum_{h=1}^{H} w_{ho} \cdot \Phi\Big(w_{1h} \cdot X_1 + w_{2h} \cdot X_2 + b_h\Big) + b_o\right) \tag{2}$$

where w_{1h} and w_{2h} are the weights of the neurons in the hidden layer; b_h is the biases of the neurons in the hidden layer; w_{ho} and b_o are the weights and bias of the output layer; H is the number of neurons in the hidden layer (10 in this study); X_1 and X_2 are the samples vectors of the two predictors; and $\Psi(\cdot)$ and $\Phi(\cdot)$ are the activation functions of the output and hidden layers, respectively. Linear and Tan-Sigmoid functions are chosen for the output and hidden layers, respectively:

$$\Psi = \frac{2}{1 + e^{-2n}} - 1 \tag{3}$$

Similar to a typical ANN structure, the NARX network consists of three layers, namely

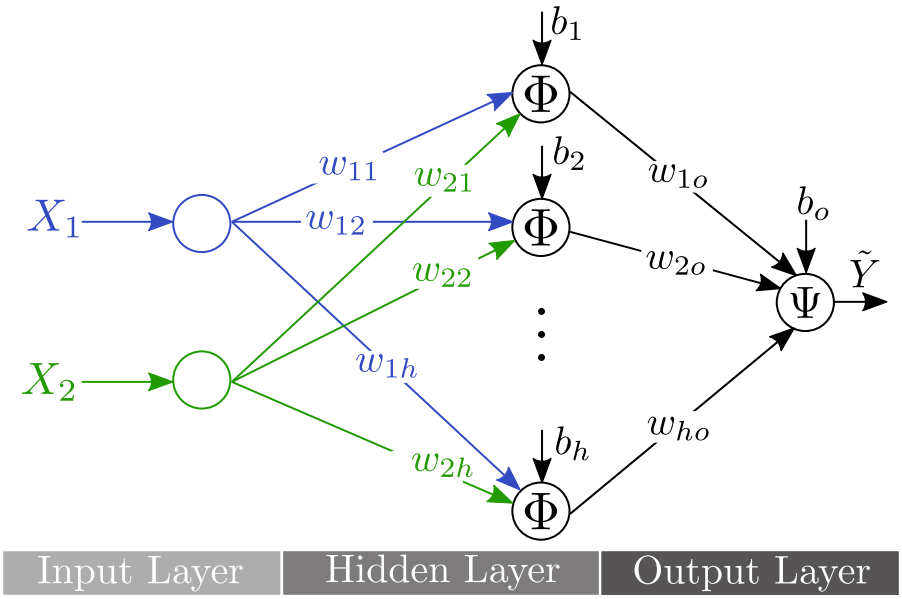


Figure 1: The structure of a feed-forward neural network model.

input, hidden and output layers. Figure 2 shows block diagram of an NARX model for estimation. Each layer consists of a specific number of neurons. In this study, a three-layer feedforward network with a sigmoid transfer function and a linear function in the hidden and output layers, respectively, is used for the approximation of the function, $f(\cdot)$. The estimated output is used as an input parameter to estimate the battery cell temperature in the future time steps. The weights and biases, shown in Figure 2, are important to optimise the network performance. Tuning the values of the weights and biases occurs during training of the NARX model. If there are any difference between estimated and actual output values, weights and biases will be adjusted during training process using a gradient descent-based (i.e., Levenberg-Marquardt in this study) algorithm to minimise the errors. Network training process will continue with the same input and output values until the error reaches to an acceptable limit. Tapped delay line (TDL), shown in Figure 2, is an embedded memory in feedforward network, which stores present and previous time series as per the defined delay.

Two different architectures of NARX are used for training and finally estimating battery cell temperature, which are series-parallel and parallel architectures. Series-parallel architecture, shown in Figure 3a and Eq.(4), is used during training stage, where actual battery cell temperature is available from the past observations. The future values are estimated from the present and past values of the input variables, $u(t)...u(t-n_u)$ and true battery cell temperature from the past, i.e., $y(t-n)$,...,$y(t-n_y)$. The advantage of this model is that the trained network has

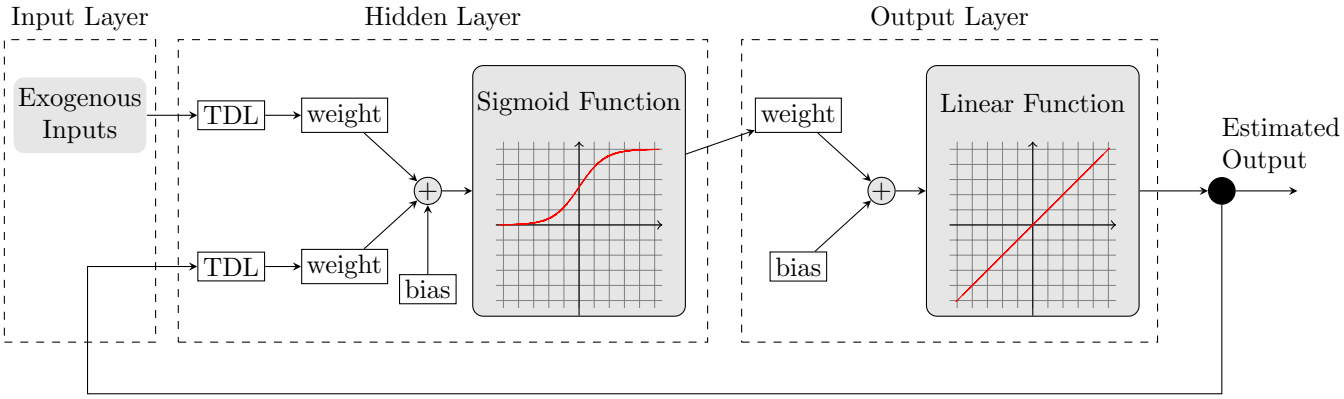


Figure 2: NARX structural model based on MLP neural network

a purely feedforward architecture and static back-propagation can be used for training. The series-parallel architecture can be represented as:

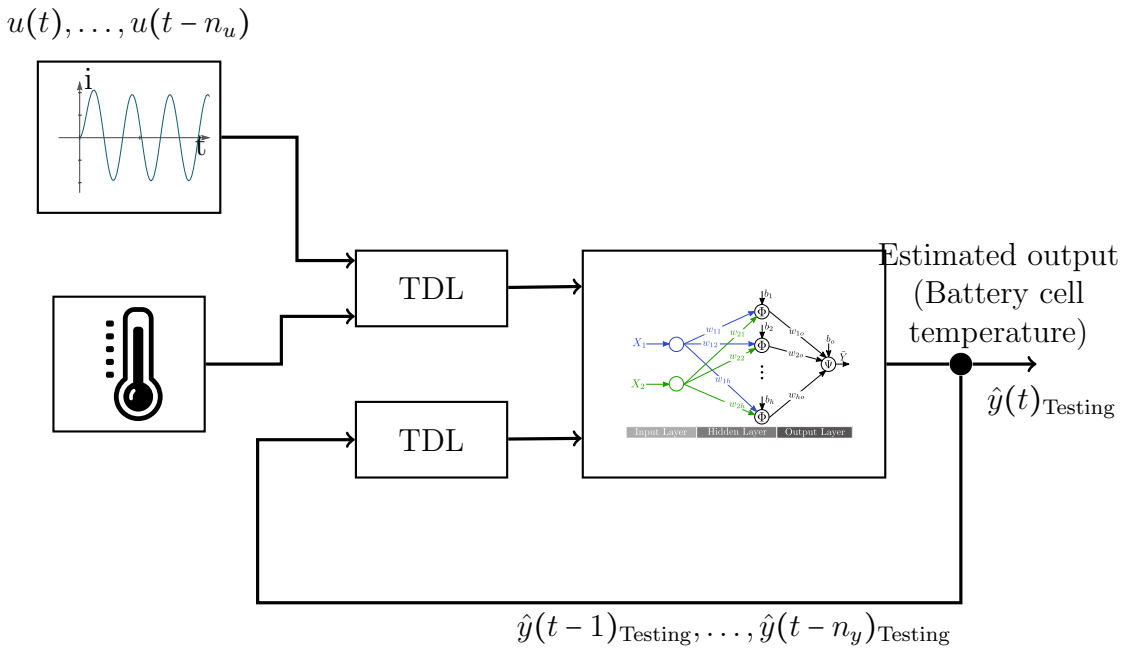
$$\hat{y}(t)_{\text{Training}} = \hat{f}(y(t-1), y(t-2), .., y(t-n_y), u(t), u(t-1), u(t-2), .., u(t-n_u)) \quad (4)$$

where, $\hat{y}(t)_{Training}$ is the Series-parallel mode output. However, the series-parallel architecture is not capable of estimating multi-step ahead cell temperature. Therefore, parallel architecture, shown in Figure 3b and Eq.(5), is used to perform future estimation for the test data and real-world simulation. The estimation is performed by taking the past and present values of the input parameters, i.e., $u(t)$,...,$u(t-n_u)$ and previously estimated battery cell temperature values, $\hat{y}(t)_{\text{Testing}}$,...,$\hat{y}(t-n_y)_{\text{Testing}}$ during testing stage, the results of which will be shown and discussed in section 4. The parallel NARX architecture can be shown mathematically by:

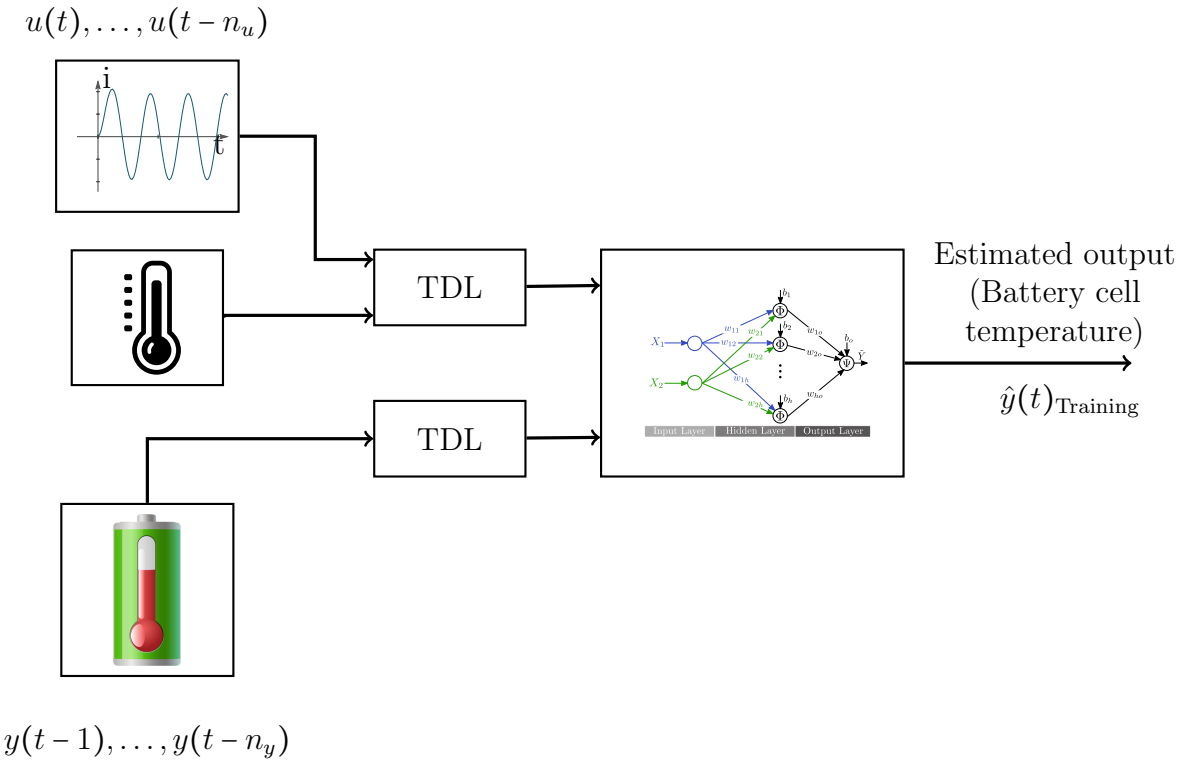
$$\hat{y}(t)_{\text{Testing}} = \hat{f}(\hat{y}(t-1)_{\text{Testing}}, \hat{y}(t-2)_{\text{Testing}}, .., \hat{y}(t-n_y)_{\text{Testing}}, u(t), u(t-1), u(t-2), .., u(t-n_u)) \quad (5)$$

where, $\hat{y}(t)_{Testing}$ is the parallel mode output.

The number of neurons in the input and output layers of the NARX model are equal to the number of independent (i.e., current and ambient temperature) and dependent variables (i.e., battery cell temperature), respectively. Only one hidden layer is considered in the NARX model as per universality theorem [29], where neural network with single hidden layer is able to approximate continuous function with desired accuracy most of the time. Selecting the number of neurons in the hidden layer is important to have the most efficient and accurate network. Employing few or too many neurons in the hidden layer will result in underfitting or overfitting, respectively. Underfitting occurs when a small number of neurons are unable to learn the patterns in the training signal adequately. On the other hand, overfitting happens when the



(a)



(b)

Figure 3: NARX Architectures: (a) Series-parallel architecture for NARX during training; (b) Parallel architecture for NARX during testing and actual simulation

trained model rely on the training dataset too much, which might provide unacceptable results for the test dataset. In addition to the importance of the number of neurons in the hidden layer, the time delay (i.e., the number of samples from the past) of the exogenous input and feedback signals are important in the performance and accuracy of the model. A sensitivity analysis has been proposed in this study to select the best number of neurons in the hidden layer and time delay in the feedback signal. In this analysis, different time delays (i.e., 10, 20, 30, 40, 50 and 60 minutes) and number of neurons (i.e., 2, 4, 6, 8, 10, 12, 14, 16, 18, 20) have been used to train the neural network and then test the model with completely different set of data. Data is normalised for training and model evaluation purposes throughout this study, where maximum value of input and output parameters in the historical data are considered as the base value for normalisation. Normalising the original data can eliminate negative influence of large numbers, which is apt to improve the convergence rate of the training process.

Three standard measures are used in this study to quantify the accuracy of the NARX model in estimating battery cell temperature. Root mean squared error (RMSE), correlation coefficient (R) and adjusted R^2 ($R^2_{Adjusted}$) given in Eq. (6), Eq. (7) and Eq. (8), respectively, are used to determine the trained network outcomes in training, testing and validation stages. The best trained NARX model is indicated by the RMSE value close to 0, R and $R^2_{Adjusted}$ value close to 1.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(T_{a_i} - T_{m_i})^2}{N}} \tag{6}$$

$$\text{R} = \frac{\sum_{i=1}^{N}\left[(T_{a_i} - \overline{T}_a) \times (T_{m_i} - \overline{T}_m)\right]}{\sqrt{\sum_{i=1}^{N}(T_{a_i} - \overline{T}_a)^2 \times \sum_{i=1}^{N}(T_{m_i} - \overline{T}_m)^2}} \tag{7}$$

$$R^2_{Adjusted} = 1 - \frac{(1 - R^2)(N - 1)}{N - k - 1} \tag{8}$$

where, in Eq. (6), T_{a_i}, T_{m_i} and N represent actual battery cell temperature, estimated battery cell temperature using NARX model and the total number of samples, respectively. In Eq. (7), $\overline{T}_a$ denotes mean value of actual temperature for battery cell. Similarly, $\overline{T}_m$ represents the average of estimated temperature. In Eq. (8), k represents the number of variables in the

model. Since network training involves solving an optimisation problem, values of biases and weights might be different in multiple runs of the training process with the same input and output values. This also can affect the performance of the network when it is used for different test dataset because different NARX network can give altered results for the same set of test samples. To resolve this issue, neural network is trained several times for the better accuracy.

In this model, one year of minute-by-minute data has been used to train a NARX network for estimating battery cell temperature. One year of data contains significant number of charging-discharging and ambient temperature data points, which covers many different combinations.

In this paper, two types of model are created. In the first type, data from the entire training set is used to develop a single NARX model for the entire year (called universal NARX model). In the second type, training data is partitioned in four different seasons and a specialised NARX model is developed for each season (called seasonal NARX model). In this case, apart from whole-year modelling structure, seasonality has been given priority since ambient temperature is highly dependent on the seasonal variations with different characteristics. Moreover, solar panels generation behaves differently in various seasons because it dominantly depends on the solar irradiation and ambient temperature. As solar radiation varies in seasons, the charging and discharging activities of the BESS varies accordingly. Therefore, more fluctuations in charging and discharging may occur during winter compared to summer or vice versa. As will be shown in the simulation results, the seasonal NARX model outperforms the universal model considerably.

3. System Under Study

In this study, operational data from a grid-connected 600 kW/760 kWh lithium-polymer battery in a 3.275 MWp PV plant is considered. The plant is located at the University of Queensland (UQ) Gatton campus, Australia. This plant is connected to the local 11 kV distribution network. Figure 4 and Table 1 show a detail overview of the battery configuration and characteristics in the plant.

The total BESS capacity and power are divided into two battery banks. Each battery bank has four racks of battery modules in parallel and each of the racks comprises ten series of battery modules. Each module has two parallel strings of 18 battery cells in series. Battery cell specification details are shown in Table 1. Please refer to [5] and [30] for more details.

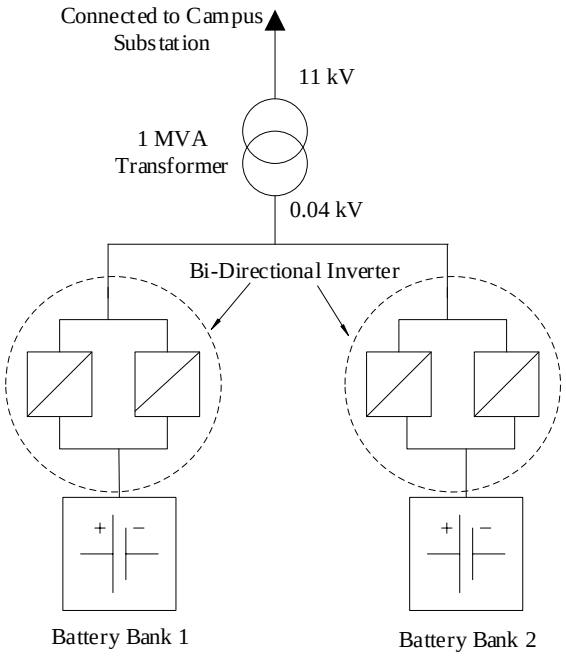


Figure 4: Schematic diagram of the BESS configuration in the PV plant

Table 1: Battery cell specifications

Description	Specification
Battery Cell Type	Lithium Polymer
Cell Capacity	75 Ah
Cell Voltage	2.7 V to 4.1 V, average 3.7 V
Cycle-Life	4000 Cycles at 80% Depth of Discharge (DoD)
Charging Temperature	10 to 35 °C
Discharging Temperature	-10 to 55 °C

In order to keep the battery cells and battery room temperature within an acceptable range, a cooling system is used in the battery container. Battery module and cell ventilation is achieved through a) passive cooling mechanisms including air vent holes along the sides of the battery casing and spacing of cells, which ensures an even temperature distribution between the cells, and b) active cooling consisting of rack fans and air-conditioning units with 7.7 kW rated cooling capacity. Three fans are placed at the top of each battery bank to draw air up through vents in the front panel by passing through the modules and out at the top. The fans start operating when cell temperature reaches 29 °C. The consumption of each rack fan is 44.64 W, which is powered by an external +24 V DC system. Battery modules are spaced inside the rack with gaps to allow airflow. The air conditioner is designed for maximum 30 °C internal ambient and 50 °C outside temperature [5, 27]

The air conditioning unit is set at the room temperature of 23 °C, which indicates that the evaporator of the air conditioning unit starts its operation along with fan when the room

temperature reaches above 23 °C. The room temperature is measured and recorded by the thermostat installed with the air conditioning unit. The higher the room temperature, the more energy consumption by the air conditioning.

A sophisticated and comprehensive measurement and data management and warehousing system is implemented in the UQ Gatton Solar Research Facility (GSRF). About 1390 points within the BESS and its container are monitored through measurement and data logging system, which gives detailed insights into the battery operation on the cell and module levels. The collected data is remotely accessible through a Wonderware Historian system interface. A significant number of variables including weather parameters, DC and AC side parameters of the inverters are measured and stored in Wanderware Historian system. The plant data logging, including BESS, is performed using a Delta Mode operation. Although data is sampled at a 1-second rate, the Delta Mode operation is adopted to reduce data storage capacity requirement. For further details, please see [5, 27]. In this study, data with 1-minute resolution (60 samples per hour) have been used for training and testing the model, which are further explained in the following section.

4. Simulation Study

In this paper, two independent parameters, i.e., battery current and ambient temperature, have been considered to create and test battery cell temperature estimation model. The following data selection and analysis are carried out in the simulation studies for training and assessing the NARX networks:

- For training purposes, 1-minute data of around 12 months from 1^{st} April 2016 to 31^{st} March 2017 have been selected, which consists of 516,960 samples (359 days excluding 6 days used of each season that is used for testing) overall. Entire year is covered in the data to represent seasonal differences in the real-world condition. A universal NARX model is created for battery cell temperature estimation for the entire year. Another 90 days equivalent data is used for validation and testing the trained model during training process in the MATLAB Toolbox. Please note that the training, validation, and testing days are selected randomly and each simulations study has been run for several times to check on the robustness of the results. Finally, the average results are reported in this section.
- Seasonality analysis has been carried out by developing separate NARX model for each

season (i.e., seasonal NARX model). Similar procedure as universal model has been followed for training each seasonal model. Multiple days of each season were removed from the training dataset as for assessing performance of the respective seasonal model.

• For further assessment of universal network model, 56 days of data from 10^{th} July 2017 to 31^{st} October 2017 as well as the test days used for seasonal model assessment are used. The impact of time delay on the performance of the model also has been analysed for feedback signal by testing the network with test days and NARX with feedback configuration, as shown in Figure 3b.

• NARX training has been repeated for different number of neurons (i.e., 2, 4, 6, 8, 10, 12, 14, 16, 18 and 20 minutes) and time delay (i.e., 10, 20, 30, 50 and 60 minutes), separately, as a sensitivity analysis to reveal their impact on the accuracy of the model and to find the best number of neurons and time delay for the rest of the simulation studies. The assessment of each model with different time delay was carried out by arbitrarily selected test data.

• A comparison study between the universal and seasonal NARX models has been carried out in the simulation study to determine the best NARX model.

MATLAB/Neural Net Time Series toolbox is used in this study for training and testing NARX models. During training process in MATLAB Neural Net Time Series toolbox, battery cell temperature is estimated for 1 minute ahead. In reality, however, we need estimated cell temperature for longer time horizons. For instance, a day-ahead EMS of a microgrid plans the system operation for 24 hours ahead. In this case, it is needed to estimate battery cell temperature for a given charge-discharge profile in the next 24 hours. To estimate temperature in multi-hour ahead, the trained model is recursively used, where the estimated cell temperature from previous time step (1 minute before) is used as an input parameter in the feedback loop. For instance, if 24 hours ahead estimation is required for a given charge-discharge and ambient temperature profiles, the estimation will be repeated 1440 times to cover the entire period.

4.1. Sensitivity analysis on the number of neurons and time delay

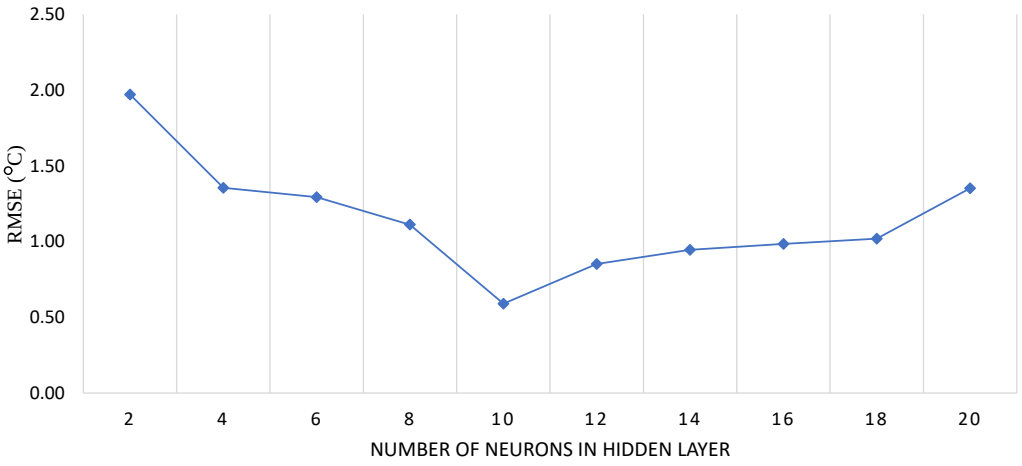
The number of neurons in the hidden layer and time delay of feedback signal are determined by sensitivity analysis. To find the appropriate number of neurons, different NARX model is trained with different number of neurons in the hidden layer, while input and output samples were remained the same. The number of neurons in the hidden layer is increased from 2 to

20 with an interval of two neurons to determine the optimal network model. Each model was run for 20 times for a specific number of neurons to reduce the impact of initialisation and optimisation of the weights and biases on the results. Average estimation error on the test dataset is used to determine the best number of neurons. As per Figure 5a, the average value of RMSE is higher when the number of neurons is at the lower and higher ranges. 10 neurons resulted in only 0.59 °C average RMSE, indicating the best performance among different cases. Therefore, NARX model with 10 neurons in the hidden layer will be used in the rest of the paper.

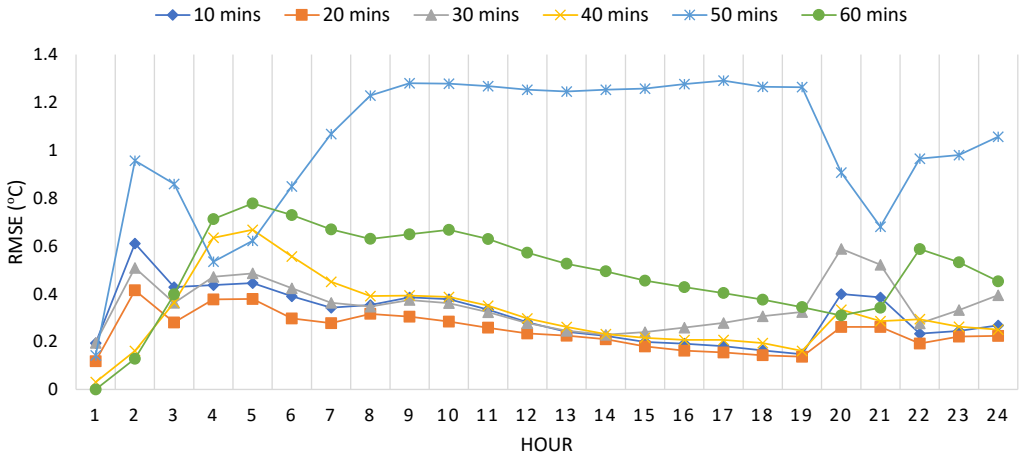
In order to find the best time delay of the feedback signal for the parallel network shown in Figure 3b, different NARX models are trained for different time delays (i.e., 10, 20, 30, 40, 50 and 60 minutes). In each case, the number of neurons in the hidden layer was 10 and the performance of the NARX model is assessed by using the test dataset. 56 days of data is randomly selected with 24 hours horizon for assessing the network performance. Hourly average RMSE values are calculated from minute-by-minute data of 56 days test data, which are shown in Figure 5b. From the figure, it is clear that 20 minutes feedback delay yields the best performance on the test dataset. 10 and 30 minutes time delay also shows reliable performance. However, creating a model with 20 minutes delay is comparatively more efficient and accurate. For the rest of the paper, 10 neurons will be considered in the hidden layer and cell temperature feedback of 20 minutes will be used. Please note that the sensitivity analyses are done for 1 minute ahead estimation.

4.2. The universal NARX model

This section presents the universal NARX model training and its performance estimation with completely different set of test data. Table 2 and Table 3 show the parameters for network training and selection of data with outcomes during training process, respectively. 449 days (excluding 6 days used for evaluating seasonality in the next section) of the data have been used for training, validation and testing. 359 days (80%) or a complete year of data has only been used to train the NARX model. 90 days (20%) of data for validation and testing stages (45 days each) are randomly chosen from the remaining data to validate the model during training stage and evaluate the trained model during testing stage. Since the data selection affects the performance of the NARX model, the training process has been repeated 10 times to



(a)



(b)

Figure 5: Sensitivity Analysis: (a) Average RMSE of different NARX model with different number of neurons in the hidden layer; (b) Sensitivity analysis results on the cell temperature feedback delay

minimise the impact of the training process and to choose the best NARX model. It has been observed that although different combinations provide different results, the performance of the model (i.e., RMSE, R and $R^2_{Adjusted}$) remains almost the same. **Validation dataset** is used to minimise overfitting during training by MATLAB toolbox. This dataset does not adjust the weights of the network model but weights are checked for accuracy in the model estimation. **Testing dataset** is completely unseen set of samples that is used by the Neural Net Time Series toolbox in MATLAB at the end of training to provide an unbiased assessment of the trained model. A good performance of a model on the test dataset is the key to a successful modelling.

Table 2: Parameters for training the NARX networks

Parameter	Description
Problem	Nonlinear auto-regressive time-series network with external input (NARX)
Training Algorithm	Levenberg-Marquardt
Input	Charging-Discharging Current and Ambient Temperature
Output	Battery Cell Temperature
Data Division	Training (80%), Validation (10%) and Testing (10%)
Hidden layer size	10 neurons
Feedback Delay	20 minutes

Table 3: Data selection and training outcomes from MATLAB Neural Net Time Series toolbox

Data Type	Number of Sample	Sample Percentage	Regression (R)	Adjusted R^2	RMSE (°C)
Training	516,960	80%	0.989	0.978	0.452
Validation	64,800	10%	0.987	0.974	0.551
Testing	64,800	10%	0.99	0.98	0.415

The performance of the universal model in comparison with the actual data is tabulated in Table. 3 for training, validation, and testing processes during training in the MATLAB Neural Net Time Series toolbox, where one time step ahead estimation is carried out. The RMSE, coefficient of determination, i.e., R, and Adjusted R^2 are calculated using Eq.(6), Eq.(7) and Eq.(8), respectively. There is a close agreement between actual battery cell temperature values with those estimated by the network. According to Table 3, the $R^2_{Adjusted}$ values for training, validation and testing are very close to 1, which indicates a good model fit. Moreover, it indicates that the model is not overfitted by providing a true goodness of fit by yielding similar performance for test dataset. In addition, the RMSE is 0.454 °C, 0.55 °C and 0.42 °C for training, validation and testing, respectively, which is very accurate for many applications such as EMS. As per both assessment criteria of the trained network, it can be stated that the universal model is able to provide a good estimation of the battery cell temperature for many

combination of input parameters.

The universal NARX model is further validated by testing with completely different set of data. In this analysis, 56 days of data (outside of training, testing and validation dataset) have been considered to estimate battery cell temperature in 24 hours ahead (minute-by-minute samples) using the trained model recursively. Figure 6 shows the hourly-averaged RMSE histogram of the 56 days of estimation, where 96.1% of incidents are experiencing an RMSE below 1 °C. Only 3.9% of incidents are yielding more than 1 °C. The majority of the incidents are between 0.17 °C and 0.22 °C RMSE range. Also, it is clear from the histogram that estimated RMSE never exceeded 1.95 °C for the 56 test days. It indicates a relatively robust performance of the NARX model, which can be effective in determining battery cell temperature in advance with a high accuracy. On average, the computation time for training the universal model and estimating the battery cell temperature for each day are 904 seconds and 40 seconds, respectively. It indicates a fast process in estimating battery cell temperature.

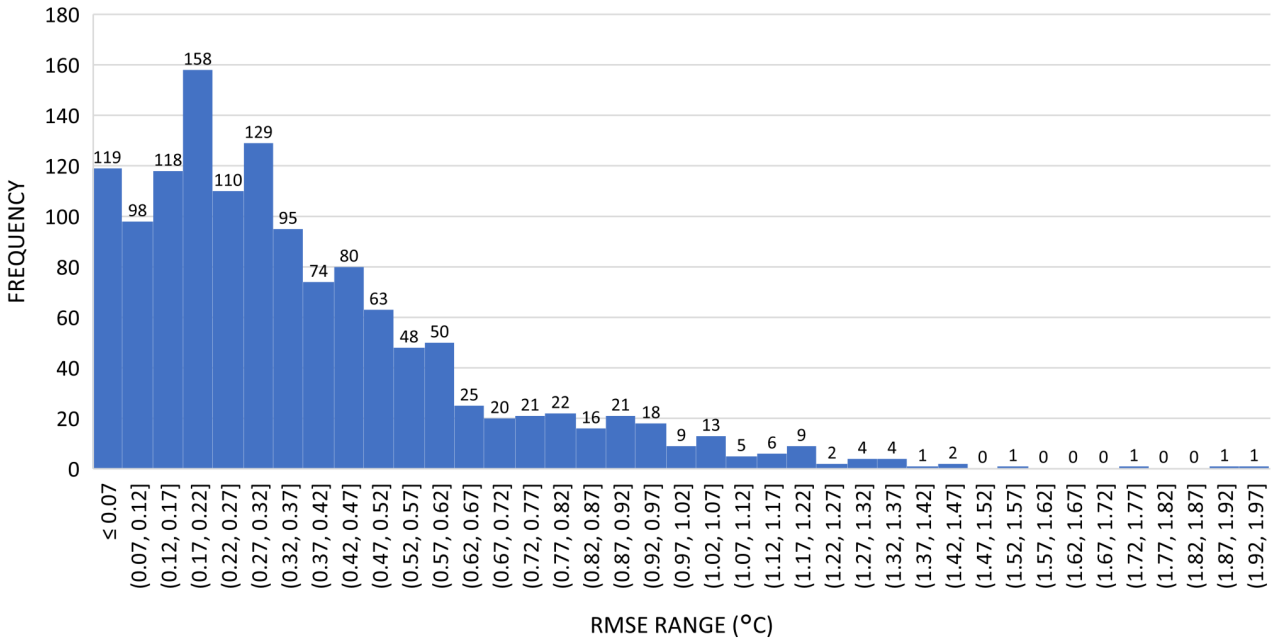


Figure 6: Hourly RMSE histogram for 56 days of test data obtained by the universal model for 24-hour ahead

4.3. Seasonal Effects

In order to evaluate the seasonal effects, an NARX model is trained for each season in this subsection, and their performance is compared with the universal model. The battery cell temperature estimation model for each season has been assessed using completely different set of data extracted from each season, explained in Table 4. The percentage of seasonal

data for training, testing, and validation differs from the the percentages in universal model, as given in Table 5. This is because there are fewer samples available for testing and validation for the seasonal model compared to the universal one. Using this approach, the number of testing samples is the same as the universal model. It is worth mentioning that the difference of outcome using 80 : 10 : 10 percentages was trivial without a significant deterioration of accuracy.

Table 4: Seasonal data selection for training and evaluation

Season	Months	Year	Days for training	Test Days for evaluation
Spring	September, October, November	2017	85	6 Days
Summer	December, January, February	2016-2017	84	6 Days
Autumn	March, April, May	2017	86	6 Days
Winter	June, July, August	2017	86	6 Days

Similar NARX configuration with 10 neurons in the hidden layer and feedback time delay of 20 minutes are used in this analysis. Randomly selected test days from each season have been extracted for assessing the trained model. In order to make fair comparison, the same six days were taken out from the training dataset in the universal model. The computation time for training each seasonal model is 50 seconds to 65 seconds, while estimating the battery cell temperature for an entire day is 40 seconds on average. This indicates a fast process in training and testing seasonal models.

Table 5: Seasonal data selection and training outcomes from MATLAB Neural Net Time Series toolbox

Season	Data Type	Number of Sample	Sample Percentage	Regression (R)	Adjusted R^2	RMSE (°C)
Spring	Training	85,680	70%	0.991	0.982	0.343
	Validation	18,360	15%	0.983	0.966	0.702
	Testing	18,360	15%	0.991	0.982	0.358
Summer	Training	84,672	70%	0.989	0.978	0.479
	Validation	18,144	15%	0.986	0.972	0.593
	Testing	18,144	15%	0.992	0.984	0.342
Autumn	Training	86,688	70%	0.991	0.982	0.286
	Validation	18,576	15%	0.993	0.986	0.276
	Testing	18,576	15%	0.99	0.98	0.31
Winter	Training	86,688	70%	0.995	0.99	0.33
	Validation	18,576	15%	0.992	0.984	0.337
	Testing	18,576	15%	0.992	0.984	0.336

Table 5 shows the number of samples and outcomes during training, validation and testing stages for each seasonal NARX model from MATLAB Neural Net Time Series toolbox. Adjusted R^2 values for each season are close to 1 in all cases, which indicates an outstanding fit to the model by drawing a close relationship between estimated and actual battery cell temperature for all four seasons. The RMSE value has also been considered to evaluate the model. According

to the respective table of each season, the RMSE values are below 0.71 °C. Most importantly the RMSE values during testing process are below 0.5 °C in all seasons, which shows a close agreement between estimated and actual battery cell temperature.

Figure 7 shows the performances of the seasonal NARX models for the test data on an hourly (i.e., RMSE hourly) basis. The battery cell temperatures are estimated recursively for 24 hours ahead, as explained at the beginning of Section 4. To obtain hourly RMSE values, average of minute-by-minute RMSEs are calculated for every hour. The values are shown for summer only in this figure due to similar performance experienced for the rest of the seasonal models. It can be observed from the figure that the RMSE value is lower than 1 °C, except a couple of hourly incidents on day 3, and the differences between days are minor. The RMSE outcomes during hour 1 to 3 and hour 10 to 16 (where battery operates rarely) of each day are comparatively better than the RMSE values during other hours. However, the differences are insignificant, which pointing out to a robust model performance throughout the whole day.

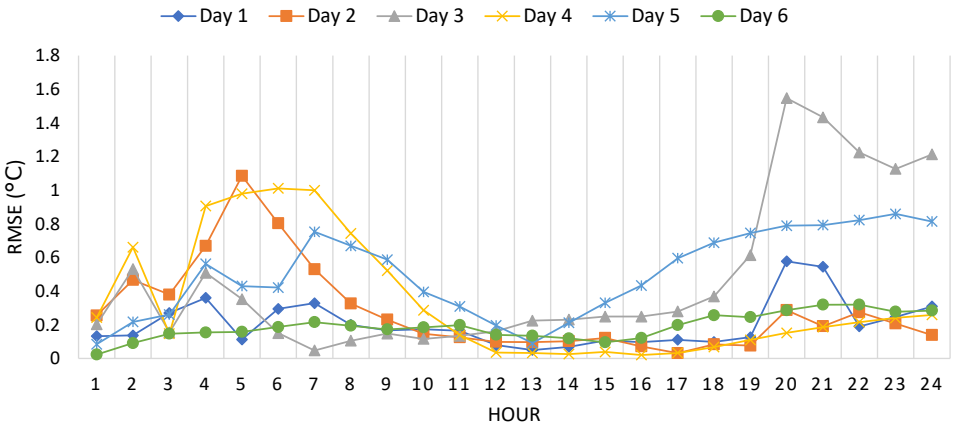


Figure 7: RMSE of the test days during summer season using recursive estimation for 24 hours ahead.

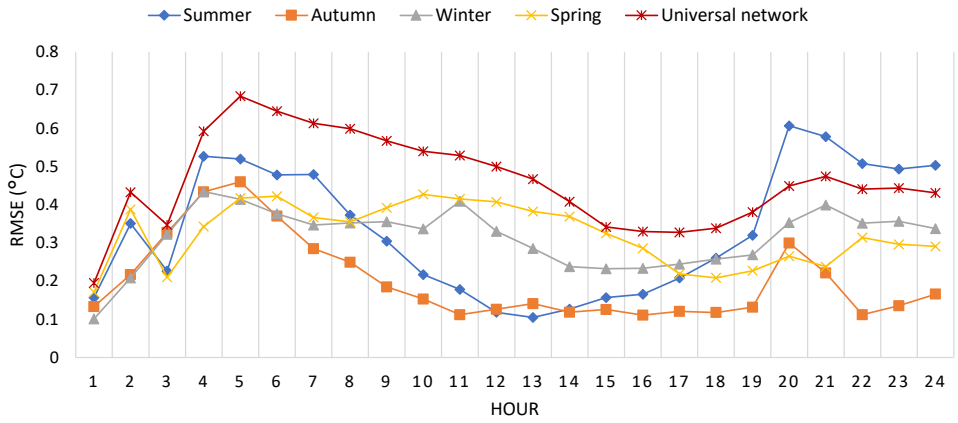


Figure 8: Seasonal comparison based on the average hourly RMSE obtained by using seasonal and universal network models.

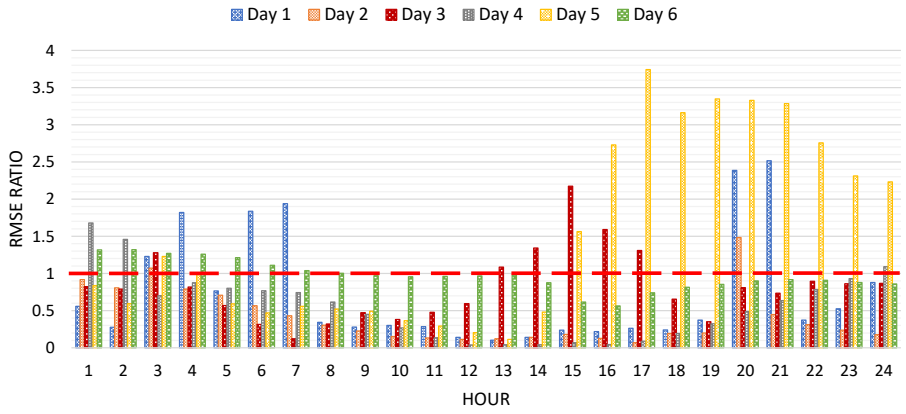
Figure 8 shows the average RMSE comparison between seasonal and universal models. In this case, the average RMSE values of the test days for each season using seasonal NARX model and 24 days (4 Seasons × 6 Days) using the universal NARX model have been considered for each hourly RMSE. It can be seen that summer and spring seasons are yielding maximum of 0.66 °C on average using the seasonal network model. On the other hand, autumn, winter and spring have RMSE values of 0.46 °C, 0.43 °C and 0.42 °C, respectively. Although couple of hours during summer show more RMSE than autumn, winter and spring, the daily average RMSE value differences between seasons are not significant, which suggests a better performance of the seasonal models in general. In comparison with the universal model, it is clear that the seasonal NARX models yield at least 0.1 °C to 0.3 °C smaller RMSE during most of the hours of a day for the same test days compared to the universal model.

In order to investigate further on the performance of the seasonal and universal NARX models, the RMSE ratio has been calculated. RMSE ratio is used in this study to compare the performance of universal and seasonal models for the similar test days. The comparison is made for every hour separately to better illustrate the variations in performance in different hours. Figure 9a shows a comparison between performance of the seasonal and universal NARX models in summer using Eq.(9). The goal is to perform a quantitative comparison between the two models. It helps to determine which model is providing better results with lower RMSE value in an hourly basis.

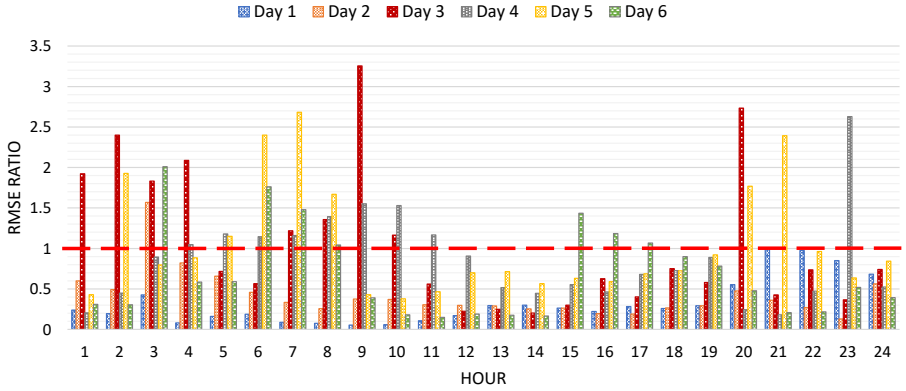
$$\textbf{RMSE Ratio} = \frac{\text{RMSE of the Seasonal Model}}{\text{RMSE of the Universal Model}} \tag{9}$$

When the RMSE ratio is lower than 1 (red dashed line in the figure), it represents a better performance of seasonal network model and above 1 always indicates better performance for the universal model. In this comparison, the same six days of test data has been used for both models and hourly estimated values up to 24 hours has been considered to compare the results.

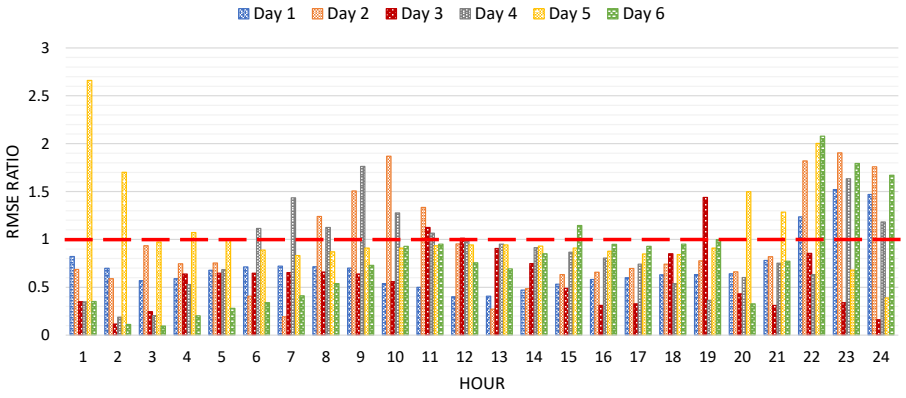
As per Figure 9a, the RMSE ratio is below 1 in most of the hourly incidents during summer test days, which denotes a better performance by the seasonal model. Although better outcomes by universal network is experienced on Day 5 for a couple of hours, the ratio is below 1 most of the time, indicating better outcomes by the seasonal network. The values are shown for summer only in this figure due to similar performance experienced for the rest of the seasonal models. Figure 9b shows almost similar outcome during autumn, where the hourly ratios are



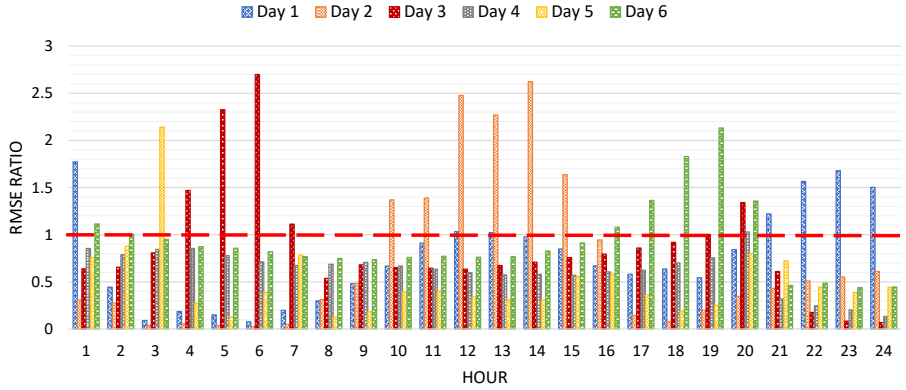
(a)



(b)



(c)



(d)

Figure 9: The RMSE ratio for the seasonal model: (a) summer; (b) autumn; (c) winter; (d) spring

less than 1 most of the time, except a couple of instances. Similar conclusion can be drawn for winter in Figure 9c. While the ratios are closer to 1 compared to summer, most of the cases are still below 1, indicating a better performance of the seasonal model compared to the universal model over the selected test days. During spring season, shown in Figure 9d, the RMSE ratio for couple of incidents from different test days are bigger than 1. However, more than 80% of the time, ratios are far below 1, which demonstrates a better performance of the seasonal network model than universal model for this season too. In few instances, hourly ratios are more than 2 in every season, which indicates the universal model is performing more than two times better than the seasonal model. However, 74.3%, 77.8%, 78.5%, 80.1% of the hours are showing better results for the seasonal model in summer, autumn, winter and spring, respectively.

Overall, it can be concluded that training the neural network with seasonal data provides a better results than the network trained with the universal data. Although there are some variations between seasons, on average, all four seasons are performing equally well and more accurate results can be obtained compared to the universal model.

Comparing the results presented in this paper against the results in [31], it can be seen that the proposed approach in this work results in lower RMSE values. The maximum difference reached to 1.48 °C for universal model. The RMSE differences between ARIMAX and NARX is slightly higher and on average the value is 2.24 °C, where ARIMAX is expressing 0.56 °C, 1.65 °C, and 1.9 °C more RMSE values than NARX models during summer, autumn, and winter, respectively. The proposed recursive estimation with NARX network in this manuscript shows better estimation results for an entire day than the ARIMAX model, which is a linear model. We have not tried ordinary NN models since they are not able to consider exogenous parameters as extra predictors in the model, which will degrade the performance of the model.

5. Conclusion

In this paper, a model based on non-linear autoregressive exogenous artificial neural network (NARX) is proposed to estimate battery cell temperature. The proposed method does not require to consider any complicated battery thermal dynamics and systematic thermo-dynamical model of the battery container. Only two dominant input parameters, namely charging-discharging current and ambient temperature, are considered for the battery cell temperature estimation,

which is shown to be sufficient. The proposed model is able to estimate battery cell temperature accurately using a relatively large historical dataset. Time-series NARX model has been chosen due to non-linear behaviour and time dependencies of the battery cell temperature. Moreover, seasonal impact on the modelling is investigated by creating different NARX model for each season. Based on the outcomes of comprehensive simulation studies, the universal model yielded a small RMSE value, where the RMSE was below 1 °C most of the times, which is quite acceptable. Compared to the universal model, the seasonal NARX models provided more accurate results by 18% to 50% on average. In addition, proposed recursive estimation with NARX network shows good estimation results for an entire day. In our future work, we will use the developed NARX model in an optimal battery operation algorithm in an EMS considering battery degradation and cooling system costs, which can improve the technical economical performance of the hybrid energy system.

Acknowledgment

This work was performed in part or in full using equipment and infrastructure funded by the Australian Federal Governments Department of Education AGL Solar PV Education Investment Fund Research Infrastructure Project. The University of Queensland is the Lead Research Organization in partnership with AGL, First Solar, and the University of New South Wales.

References

[1] S. Pourmousavi and T. K. Saha, "Evaluation of the battery operation in ramp-rate control mode within a PV plant: A case study," *Solar Energy*, vol. 166, pp. 242–254, 2018.

[2] B. Scrosati and J. Garche, "Lithium batteries: Status, prospects and future," *Journal of Power Sources*, vol. 195, no. 9, pp. 2419–2430, 2010.

[3] J. D. Bishop, C. J. Axon, D. Bonilla, M. Tran, D. Banister, and M. D. McCulloch, "Evaluating the impact of V2G services on the degradation of batteries in PHEV and EV," *Applied energy*, vol. 111, pp. 206–218, 2013.

[4] A. Millner, "Modeling lithium ion battery degradation in electric vehicles," in *Innovative Technologies for an Efficient and Reliable Electricity Supply (CITRES), 2010 IEEE Conference on*, pp. 349–356, IEEE, 2010.

[5] M. M. Hasan, S. A. Pourmousavi, F. Bai, and T. K. Saha, "The impact of temperature on battery degradation for large-scale bess in a PV plant," in *Universities Power Engineering Conference (AUPEC), 2017 Australasian*, pp. 1–6, IEEE, 2017.

[6] Q. Wang, B. Jiang, B. Li, and Y. Yan, "A critical review of thermal management models and solutions of lithium-ion batteries for the development of pure electric vehicles," *Renewable and Sustainable Energy Reviews*, vol. 64, pp. 106–128, 2016.

[7] O. Gross and S. Clark, "Optimizing electric vehicle battery life through battery thermal management," *SAE International Journal of Engines*, vol. 4, no. 1, pp. 1928–1943, 2011.

[8] F. Leng, C. M. Tan, and M. Pecht, "Effect of temperature on the aging rate of li ion battery operating above room temperature," *Scientific reports*, vol. 5, p. 12967, 2015.

[9] T. M. Bandhauer, S. Garimella, and T. F. Fuller, "A critical review of thermal issues in lithium-ion batteries," *Journal of the Electrochemical Society*, vol. 158, no. 3, pp. R1–R25, 2011.

[10] J. Newman and K. E. Thomas-Alyea, *Electrochemical systems*. John Wiley & Sons, 2012.

[11] T. F. Fuller, T. Bandhauer, and S. Garimella, "Electrochemical-thermal modeling and microscale phase change for passive internal thermal management of lithium ion batteries.," tech. rep., Sandia National Laboratories, 2012.

[12] J. Liang, Y. Gan, W. Song, M. Tan, and Y. Li, "Thermal–electrochemical simulation of electrochemical characteristics and temperature difference for a battery module under two-stage fast charging," *Journal of Energy Storage*, vol. 29, p. 101307, 2020.

[13] S. Panchal, I. Dincer, M. Agelin-Chaab, R. Fraser, and M. Fowler, "Experimental temperature distributions in a prismatic lithium-ion battery at varying conditions," *International Communications in Heat and Mass Transfer*, vol. 71, pp. 35–43, 2016.

[14] L. Fan, J. Khodadadi, and A. Pesaran, "A parametric study on thermal management of an air-cooled lithium-ion battery module for plug-in hybrid electric vehicles," *Journal of Power Sources*, vol. 238, pp. 301–312, 2013.

[15] L. Lu, X. Han, J. Li, J. Hua, and M. Ouyang, "A review on the key issues for lithium-ion battery management in electric vehicles," *Journal of power sources*, vol. 226, pp. 272–288, 2013.

[16] S. Chen, C. Wan, and Y. Wang, "Thermal analysis of lithium-ion batteries," *Journal of power sources*, vol. 140, no. 1, pp. 111–124, 2005.

[17] V. H. Johnson, A. A. Pesaran, and T. Sack, *Temperature-dependent battery models for high-power lithium-ion batteries*. National Renewable Energy Laboratory City of Golden, 2001.

[18] D. Kang, P.-Y. Lee, K. Yoo, and J. Kim, "Internal thermal network model-based inner temperature distribution of high-power lithium-ion battery packs with different shapes for thermal management," *Journal of Energy Storage*, vol. 27, p. 101017, 2020.

[19] S. Basu, K. S. Hariharan, S. M. Kolake, T. Song, D. K. Sohn, and T. Yeo, "Coupled electrochemical thermal modelling of a novel li-ion battery pack thermal management system," *Applied Energy*, vol. 181, pp. 1–13, 2016.

[20] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks:: The state of the art," *International journal of forecasting*, vol. 14, no. 1, pp. 35–62, 1998.

[21] M. Castilla, J. Álvarez, M. Ortega, and M. Arahal, "Neural network and polynomial approximated thermal comfort models for hvac systems," *Building and Environment*, vol. 59, pp. 107–115, 2013.

[22] X. Wang, X. Wei, Q. Chen, J. Zhu, and H. Dai, "Lithium-ion battery temperature on-line estimation based on fast impedance calculation," *Journal of Energy Storage*, vol. 26, p. 100952, 2019.

[23] M. Lucu, E. Martinez-Laserna, I. Gandiaga, K. Liu, H. Camblong, W. Widanage, and J. Marco, "Data-driven nonparametric li-ion battery ageing model aiming at learning

from real operation data-part b: Cycling operation," *Journal of Energy Storage*, vol. 30, p. 101410, 2020.

[24] I. A. Basheer and M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application," *Journal of microbiological methods*, vol. 43, no. 1, pp. 3–31, 2000.

[25] S. Chen and S. Billings, "Neural networks for nonlinear dynamic system modelling and identification," *International journal of control*, vol. 56, no. 2, pp. 319–346, 1992.

[26] A. Mellit, S. Kalogirou, S. Shaari, H. Salhi, and A. H. Arab, "Methodology for predicting sequences of mean monthly clearness index and daily solar radiation data in remote areas: Application for sizing a stand-alone pv system," *Renewable Energy*, vol. 33, no. 7, pp. 1570–1590, 2008.

[27] M. M. Hasan, "Thermal modelling and cost analysis for large-scale battery energy storage system (BESS) in grid-connected pv plant," Master's thesis, School of Information Technology and Electrical Engineering, 8 2019. `https://espace.library.uq.edu.au/view/UQ:5b302b9`.

[28] L. Ljung, "Black-box models from input-output measurements," in *Instrumentation and Measurement Technology Conference, 2001. IMTC 2001. Proceedings of the 18th IEEE*, vol. 1, pp. 138–146, IEEE, 2001.

[29] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural networks*, vol. 4, no. 2, pp. 251–257, 1991.

[30] M. Alam, R. Yan, T. Saha, A. Chidurala, and D. Eghbal, "Learning from a 3.275 mw utility scale pv plant project," in *Proceedings of the CIGRE Paris Conference*, 2016.

[31] M. M. Hasan and S. Ali Pourmousavi, "Battery cell temperature estimation model and cost analysis of a grid-connected pv-bess plant," in *2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, pp. 1804–1809, 2019.

Response to Decision Letter

We thank the Editor and Reviewers for their thought-intriguing comments and useful suggestions. We respond to each of them by both providing clarifications in this letter and by modifying the paper accordingly. The reviewers' comments are shown in blue fonts in this document, and the revised texts in the paper are highlighted in red fonts.

Reviewers' Comments

Reviewer 1

The manuscript reported an approach to estimate cell temperature using NARX neural network for battery energy storage system. The paper is well-written in language and easy to understand. The overall organization of sections and flow logic are acceptable. Most of the figures and their captions present the idea clearly. The NARX neural model is described clearly, and two models are established as universal model and seasonal model, which is thoughtful and specific work. This work is significant to the thermal management for energy storage system. However, there are some questions and incomprehension in this manuscript, as follows.

Comment 1.1: Line 257, 1-minute data has been considered for training and testing the model. How to explain the 1-minute data? Every sample includes 1-minute data in each day? If so, in the simulation, only 1-minute data was considered for training and testing, so whether is the data representative?

Answer: We would like to thank the Reviewer for her/his comments. Apologies for the confusion. The "1-minute" refers to the resolution of the data, i.e., there are 60 samples for every hour. We have modified page 14, line 277 of the text to remove the ambiguity.

Comment 1.2: In Table 2, the percentage of Training, Validation, Testing is 8:1:1 in universal model, and in the seasonal model, the percentage is 7:1.5:1.5. Why? and whether it affect the comparison between the universal and seasonal model?

Answer: The percentage of seasonal data for training, testing, and validation differs from that of the universal model since there are fewer sample points available for testing and validation of the seasonal model compared to the universal model. Therefore, we decided to reserve more samples for testing. Nevertheless, the difference of outcome using similar

percentages was trivial and it does not affect the estimation performances while testing the model with completely separated test data/days. It is worth mentioning that similar test days are used to check the universal and seasonal trained models for fair comparison. To clarify this, we have modified Section 4.3, page 19, line 402.

Comment 1.3: In this work, the BESS has an active cooling systems 7.7 kW. How does the active cooling systems operate? It works in whole charging-discharing process, or other ways. For other BESS, it has different active cooling systems, are this models available for the temperature estimation?

Answer: In order to keep the battery cells and battery room temperature within an acceptable range, a cooling system is used in the battery container. Battery module and cell ventilation is achieved through a) passive cooling mechanisms including air vent holes along the sides of the battery casing and spacing of cells, which ensures an even temperature distribution between the cells, and b) active cooling consisting of rack fans and air-conditioning units with 7.7 kW rated cooling capacity. Three fans are placed at the top of each battery bank to draw air up through vents in the front panel by passing through the modules and out at the top. The fans start operating when cell temperature reaches 29 °C. The consumption of each rack fan is 44.64 W, which is powered by an external +24 V DC system. Battery modules are spaced inside the rack with gaps to allow airflow. The air conditioner is designed for maximum 30 °C internal ambient and 50 °C outside temperature.

The air conditioning unit is set at the room temperature of 23 °C, which indicates that the evaporator of the air conditioning unit starts its operation along with fan when the room temperature reaches above 23 °C. The room temperature is measured and recorded by the thermostat installed with the air conditioning unit. The higher the room temperature, the more energy consumption by the air conditioning.

Section 3, page 13, line 261, of the paper is modified to reflect this discussion.

Reviewer 2

This paper presents a data-driven model based on a Non-linear Autoregressive Exogenous neural network to estimate battery cell temperatures in a utility-scale BESS, considering strongly-correlated independent variables such as charging-discharging current and ambient

temperature.

Comment 2.1: The Non-linear Autoregressive Exogenous neural network is not new in literature. What are the main contributions made by the paper? It is not very clearly described.

Answer: We would like to thank the reviewer for her/his comments. We agree with the reviewer that NARX modelling tool has been used in other literature for different purposes. However, to the best of our knowledge, the NARX application for day-ahead large-scale battery temperature estimation considering active and passive cooling mechanisms has not been reported in the literature. Ordinary neural networks have been used in the past for different purposes, however, the ordinary neural networks are not able to model temporal information in a time series. In other words, not only the input features contain useful information for estimation, but also the changes in input/output over time can provide additional information to increase the accuracy of estimation. An ordinary NN treats every sample individually; thus, it fails to model temporal information.

We added a paragraph in the Introduction section, page 5, line 97, that outlines the contributions of this work.

Comment 2.2: Pls be clear on ambient temperature. Are you referring to surface temperate of the cell?

Answer: We apologies for the confusion. Weather temperature or outside of the container temperature is referred to as ambient temperature in this study. Although, charge-discharge activities of the battery are the main reasons of battery temperature, ambient temperature is also strongly responsible for the increase in the battery cell temperature [5]. Fig. 10 in the revised manuscript illustrates the concept.

Following the Reviewer's comment, we have clarified this matter in Section 2, page 6, line 130, of the manuscript.

Comment 2.3: Also where and how did you measure the cell temperature?

Answer: The UQ Gatton Solar Research Facility (UQ GSRF) is a multi-million dollar plant that is designed primarily as a research facility with sophisticated BMS for each battery bank, delivered by Kokam company. It is designed to provide critical information such as cell temperature, voltage, current, power etc. About 1390 points within the BESS is monitored through measurement and data logging systems, which gives detailed insights into the battery

operation on the cell and module levels.

Following the Reviewer's comment, we have added a paragraph to explain measurement and data collection systems at UQ GSRF in Section 3, page 14, line 277, of the manuscript.

Comment 2.4: How do you collect the experimental data for the training?

Answer: Central Supervisory System (CSS) PLC receives all the data from measurement units and sensors within the BESS through a sophisticated SCADA system. All the CSS collected data for the plant is remotely accessible through a Wonderware Historian system interface. A significant number of variables including weather parameters, cell level voltage, current, and temperature, and battery inverters' DC and AC side parameters are measured and stored in Wanderware Historian system. The plant data logging, including BESS, is performed using a Delta Mode operation. Although data is sampled at a 1-second rate, the Delta Mode operation is adopted to reduce data storage capacity requirement. In this procedure, a measurement is recorded when the value is different from the previous measurement by a certain preset threshold.

Following the comment of the Reviewer, we have added these explanations to the manuscript in Section 3, page 14, line 277.

Comment 2.5: What do you mean by "...90 days (20%) of data for validation and testing stages..".

Answer: During training stage, 20% of data or 90 days equivalent data is being used for validation and testing purposes. The validation data is used during NARX training process to prevent from over-fitting. The testing portion of data is used to finally verify the robustness and universality of the trained model. In this paper, data is separated on a daily basis (24 hours window) to capture the operational and ambient temperature condition during each day and the 90 days are randomly selected.

Comment 2.6: Do you do cross-validation? 5-fold/10-fold?

Answer: For cross-validation, we ran the simulation studies by randomly selecting the days for each category, i.e., training, validation, and testing. Then, we ran simulation studies several times. The simulation results showed no significant variations in the performance of the universal and seasonal models. Finally, we reported the set of results with average performance as a measure of fairness. We added the explanations in the paper for the future audience.

Following the comment of the Reviewer, we have added a brief explanation to the manuscript in Section 4, page 14, line 300.

Comment 2.7: The full hyperparameters values should be given.

Answer: The hyperparameter values are reported in Table 2: "Parameters for training the NARX networks", and more details are provided in Section 4 (Section 4.1, 4.2, 4.3) of the manuscript.

Comment 2.8: What is the purpose of establishing the RMSE ratio in (9) where the aim is to estimate the battery cell temperature?

Answer: RMSE ratio is used in this study to show the comparison between performance of universal and seasonal models for the similar test days. Using the RMSE ratio, we can compare the performance of the seasonal and universal models. Ultimately, the model with lower RMSE value is the one with a better performance.

The manuscript has been modified on page 22, line 445 to reflect this discussion.

Comment 2.9: Can the proposed recursive estimation with NARX network shows good estimation results for different day and time of a day? Pls show the results.

Answer: The accuracy of the models are not identical in different hours and days, as shown in Fig. 7 for different test days and hours in summer, and in Fig. 8 for seasonal comparison. Nevertheless, the estimation error is mostly below $1 \circ C$ for different cases and hours. As mentioned previously in response to the Reviewer's comment, we ran the simulation studies several times by randomly selecting training, validation and testing days. no significant variations have been observed in this study, which proves the robustness of the proposed model.

Comment 2.10: There are lack of comparisons with other NN. It is not very convincing that the proposed method is the best. Not forgetting the proposed method is not new.

Answer: Thank you for your comment. The authors have published another paper, *Md Mehedi Hasan, S. Ali Pourmousavi, and Tapan K. Saha, "Battery cell temperature estimation model and cost analysis of a grid-connected PV-BESS plant", IEEE PES ISGT Asia, Chengdu, China, May 21-24, 2019*, where Autoregressive Integrated Moving Average with eXogenous inputs (ARIMAX) model was applied on the dataset. Comparing the results presented in this manuscript against the results of the previously-published paper, it can be seen that the proposed approach in this manuscript results in lower RMSE. The maximum hourly RMSE

using ARIMAX was 1.48 °C higher than the maximum hourly RMSE using NARX for the universal model. For the seasonal models, the hourly RMSE using ARIMAX was 0.56 °C, 1.65 °C and 1.9 °C more than NARX models during summer, autumn, and winter, respectively. Therefore, the proposed recursive estimation with NARX network yields much better estimation in comparison with the ARIMAX model, which is not surprising considering the linear nature of ARIMAX model and the ability of NARX in modeling non-linear relationships.

As was explained in response to the first comment by the Reviewer, we have not tried ordinary NN models because they are not able to model temporal information in time series, which will degrade the performance of the model. This is explained in the revised manuscript on page 5, line 97.

Comment 2.11: How robust is the proposed NARX as compared to other approaches based using NN?

Answer: We would like to refer the Reviewer to our response to the previous comment. We have also offered a comparison between the proposed method with ARIMAX, a widely used approach.

Comment 2.12: What is the computational time for just training?

Answer: On average computation time for training is:

- Universal model training: 15 minutes 4 seconds;
- Seasonal model training: 50 seconds to 1 minute 5 seconds

We have added these information to the manuscript in Section 4.1 on page 19, line 395 and Section 4.2 on page 20, line 412.

Comment 2.13: What is the computational time for estimating the battery cell temperature?

Answer: Training the model is the most time consuming part, while testing is fast. Computational time for estimating battery cell temperature for an entire day is 40 seconds on average. Explanations are added in Section 4.1 on page 19, line 395 and Section 4.2 on page 20, line 412 to indicate the estimation time.