Journal of Innovative Technology, Vol. 3, No. 1, pp. 57-63, March 2021
http://doi.org/10.29424/JIT.202103_3(1).0008

Forecasting Annual Electricity Consumption in Indonesia: A GRNN-based Approach

Ulin Nuha [1] and Kuang-Hui Chi [2*]

ABSTRACT

Forecasting electricity consumption has played a critical role in planning a country's electrical energy system. This paper presents a method of forecasting such annual consumption in Indonesia by the GRNN (generalized regression neural network) with k-fold cross-validation plus golden-section search and interpolation techniques. We proposed to resolve the best spread parameter which could govern the performance of GRNNs. Cubic spline and polynomial regression were performed as counterparts to be compared. Allowing for real data of more than two decades, numerical analysis indicates that the proposed method could outperform counterpart schemes, maintaining the least error of estimation of 1.47%. The well-trained model shall then be utilized to forecast Indonesia's annual electricity consumption for the next three years 2020 to 2022. Our model suggests that electricity consumption tends to increase steadily.

Keywords: Forecast, electricity consumption, GRNN, k-fold cross-validation, numerical method.

1. INTRODUCTION

Forecasting electricity demand is of great importance when it comes to planning the strategic policies of energy supply for an area in the future. This is one of the most critical measures taken by each of the governments and countries worldwide. Take Indonesia as an example, the Ministry of Energy and Mineral Resources indicated that fossil energy remains the dominant source, accounting for over 85.3 percent of the total national installed capacity while producing 35,951 GWh (Gigawatt-hours) surplus energy in 2019. Burning fossil fuels to produce electricity not only releases heat-trapping greenhouse gases and air pollutants but also yields a two-thirds loss of energy (vented as heat) at most of the power plants. In, thus focusing on this, such an energy supply-demand mismatch deserves attention, and so it is vital to develop forecasting means for the future demand as such based on the current trend in order to better avoid waste as well as reconcile the global objectives of carbon offset.

The use of electrical energy can possibly represent a development index of a certain area (Wu *et al*. 2018). Data on the issue of electricity consumption reflects potential demands for electricity production and consumption patterns, which can further enable the energy management authorities to make decisions on critical policies, such as the area of unit load, system design and management of power plants, and sales of electrical energy (Fan *et al*. 2020).

Manuscript received December 29, 2020; revised February 2, 2021; accepted February 16, 2021.

[1] Master Student, Department of Electrical Engineering, National Yunlin University of Science and Technology, Douliu,Yunlin, Taiwan 640301, R.O.C.

[2*] Professor (corresponding author), Department of Electrical Engineering, National Yunlin University of Science and Technology, Yunlin, Taiwan 640301, R.O.C. (e-mail: chikh@yuntech. edu.tw).

Forecasting electricity consumption can generally fall into the following three categories: Short, medium, and long. Short-term forecasting is used at intervals over a period of less than one week. Medium-term forecasting is typically used on a one-year basis. Long-term forecasting is for a span of more than one year (Filik *et al*. 2011). It is worth mentioning that this study is concerned with long-term forecasting.

Although there has been plenty of studies on forecasting electricity consumption, observed schemes have demonstrated accuracy, but this may appear liable to *instability*. Here, instability refers to the fact that estimates cannot converge to real data over time. Besides, previous schemes generally require a large set of samples with costly computational time, thereby causing problems, when the dataset is incomplete (Wang *et al*. 2018).

It is observed that fuzzy logic control and neural network control are the two significant categories of artificial intelligence, especially to develop approaches required to make effective decisions. Fuzzy logic control makes decisions based on the ambiguous input data, and yet this has a difficulty in determining fuzzy sets and fuzzy rules in this study. In contrast, we set out to bring forward a neural network-based scheme whose interconnected multilayer perceptron is trained adaptively to produce outputs rather than pre-programmed. As a remedy, this paper aims to propose a means to forecast Indonesia's annual electricity consumption based on a GRNN (generalized regression neural network). GRNN is considered excellent due to the fact that it requires only smaller datasets.

The neural network model solely relies on training and testing data to become practically reliable. In this way, it must have proper training data. Given the same data sources, if the selection of training data is diverse, the accuracy will also be diverse. To prevent an undue selection of training data and aggravating inaccuracy, *k*-fold cross-validation was adopted in our architecture to avoid the purpose of less biased or less optimistic estimates of electricity consumption. Meanwhile, we formulated an optimization problem of finding a key parameter for our GRNN through which the best performance could be achieved.

As far as the GRNN adopted is concerned, techniques like a genetic algorithm, particle swarm optimization, or fruit fly optimization can be applied. However, the former two take a relatively long time for the training process to occur (Polat and Yildirim, 2008; Sun *et al*. 2014), while, in contrast, fruit fly optimization in a complex environment may be trapped into local extrema (Zhang *et al*. 2019). Herein, we adopted golden-section search and parabolic interpolation instead. Note that golden-section search finds an optimum solution, while parabolic interpolation enables faster convergence (Renk *et al*. 2009). We also compared the proposed GRNN with two well-known numerical methods for performance evaluation.

The remainder of this paper is organized as follows. The next section gives a background on our methodology. Section 3 discusses and compares performance results derived from the proposed method with those from other schemes. Lastly, conclusions of this study are drawn in Section 4.

2. BACKGROUND

2.1 Generalized Regression Neural Network

GRNN, known as a memory-based network paradigm, was proposed by Specht to forecast continuous variables (Specht, 1991). GRNN is apt at addressing nonlinearity problems based on the estimation of a probability distribution function. One critical advantage of the GRNN is that its simple model can possibly make accurate predictions, using a small amount of input data. GRNN embeds a learning algorithm that evaluates the relationship between the target variable and the input variable (Zhou *et al*. 2014).

GRNN primarily consists of four layers (Fig. 1): Input layer, pattern layer, summation layer, and output layer. The input layer informs and collects data required for the next layer. The pattern layer is mainly used to classify the training operation (Heydari *et al.* 2019). The third layer deals with a summation of activation function values and a summation of training output and activation function values. The final layer, otherwise known as the output layer, handles the weighted average of all observed cases (Yip *et al.* 2014). A GRNN operates in the following lines:

1. Input layer: The number of neurons is the same as the input vector dimension in the learning sample. Each of the neurons in this layer is a simple distribution unit that directly forwards input variables to the pattern layer. The input is $X = [x_1, x_2, \ldots, x_n]$, where n denotes the number of input variables.
2. Pattern layer: Each of the neurons in the second layer computes the mapping between input and output vectors of a pattern based on Euclidean distances from neighboring points and the spread parameter (smoothing factor) σ, as expressed in Eq. (1).

$$P_i = \exp\left[-\frac{(X - X_i)^T (X - X_i)}{2\sigma^2}\right] \qquad (1)$$

where X represents the input variable, Xi is the learning sample of the i-th input neuron, $1 \le i \le n$ (Ruiming and Shijie 2020). P_i is the output of each neuron in the pattern layer. The number of neurons in the second layer is the same as that of the first layer.

Summation layer: This layer consists of simple summation and weighted summation. Simple summation adds up the output of all neurons in the second layer, as given by Eq. (2).

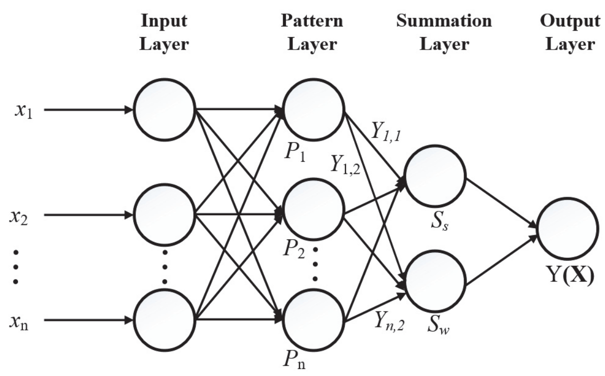

Fig. 1 GRNN Architecture

$$S_s = \sum_{i=1}^{n} P_i \qquad (2)$$

Weighted summation adds up both the connection weights and each neuron in the pattern layer, as shown in Eq. (3).

$$S_w = \sum_{i=1}^{n} y_i P_i \qquad (3)$$

where y_i represents the i-th connection weight from the second-layer neuron to the next layer neuron (Cross *et al.* 2018).

Output layer: The forecast result of the target variable can be obtained by taking the weighted summation S_w divided by the simple summation S_s, as expressed in Eq. (4).

$$Y(\mathbf{X}) = \frac{S_w}{S_s} \qquad (4)$$

Y(X) takes on the ratio of S_w to S_s derived from the third layer.

2.2 K-Fold Cross-Validation

Cross-validation (CV) is an effective technique in assessing how accurately a model will work. As Fig. 2 depicts, its basic idea is to divide the dataset, once or several times, to estimate the fallacy of each model. The training set is used to train every model, and the rest (validation set) is used to estimate the error of each model. By applying cross-validation, the best model can be obtained and so employed to predict new data with the smallest error of estimation (Arlot and Celisse 2010).

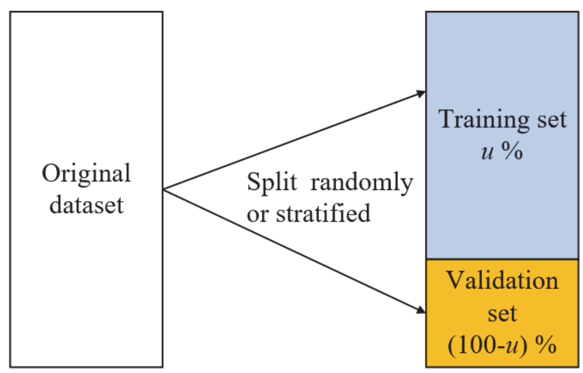

Fig. 2 Cross-validation, where *u*% of data is for training and the remaining (100-*u*)% for test

There are a number of types of cross-validation, such as *k*-fold cross-validation, repeated random sub-sampling validation, $k \times 2$ cross-validation and leave-one-out cross-validation (Rohani *et al.* 2018). Among others, *k*-fold cross-validation is usually statistically better (Barrow and Crone, 2016). It is also observed that 10-fold cross-validation tends to give better estimations than other fold configurations (Jung *et al.* 2020), so it was adopted in this study.

2.3 Polynomial Regression

The polynomial is often used for detecting curves. Polynomial regression can generally be built with a least squares model. This method can find the best curve for a group of points by means of minimizing the sum of the least square errors of the data points from the curve (Wan *et al.* 2016).

To forecast the values of interest in relation to power consumption, let us assume a system of linear equations for regression, as expressed in Eq. (5).

$$\tilde{y} = k_0\tilde{x}_0 + k_1\tilde{x}_1 + \cdots + k_i\tilde{x}_i + \varepsilon \tag{5}$$

where $\tilde{x}_0, \tilde{x}_1, \cdots, \tilde{x}_i$ are $i + 1$ basis functions (*i.e.* annual electricity consumption data). This equation can be expressed in matrix form, as given in Eq. (6).

$$\{\tilde{y}\} = [\tilde{X}]\{k\} + \{\varepsilon\} \tag{6}$$

where $\{\tilde{y}\}$ is a vector of forecast values indicating electricity consumption. $\{\tilde{X}\}$ is the matrix of values calculated by the basis functions at the measured values of the independent variable (in this case, a time-series of annual electricity consumption data). $\{k\}$ is a vector that includes the regression coefficients, and $\{k\}$ is the vector of residuals.

$$\{\tilde{y}\}^T = [\tilde{y}_1 \quad \tilde{y}_2 \quad \cdots \quad \tilde{y}_j]$$

$$[\tilde{X}] = \begin{bmatrix} \tilde{x}_{01} & \tilde{x}_{11} & \cdots & \tilde{x}_{i1} \\ \tilde{x}_{02} & \tilde{x}_{12} & \cdots & \tilde{x}_{i2} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{x}_{0j} & \tilde{x}_{1j} & \cdots & \tilde{x}_{ij} \end{bmatrix}$$

$$\{k\}^T = [k_0 \quad k_1 \quad \cdots \quad k_i]$$

$$\{\varepsilon\}^T = [\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_j]$$

where i is the number of elements in a sequence for data forecasting, and j is the number of data points.

The sum of the squares of residuals (namely vector ε) can be minimized by performing its partial derivative to each of the coefficients and adjusting the resulting equation equal to zero. Then, the outcome of Eq. (6) is a normal equation that can be expressed in Eqs. (7) and (8).

$$[\tilde{X}]^T[\tilde{X}]\{k\} = [\tilde{X}]^T\{\tilde{y}\} \tag{7}$$

$$\{k\} = \left[[\tilde{X}]^T[\tilde{X}]\right]^{-1}[\tilde{X}]^T\{\tilde{y}\} \tag{8}$$

Once the coefficient vector $\{k\}$ is resolved, the polynomial function for linear regression can be determined from Eq. (5).

2.4 Cubic Spline Interpolation

The spline interpolation normally involves sections of polynomials on a subinterval that is knitted with certain continuity conditions. A spline having the order of three or a cubic spline is three times continuously differentiable. That is, it has continuous self, first and second derivatives at the knot point. Such an interpolation technique has the utility of finding a polynomial of order three in every interval among knots. For n data points $(\tilde{x}_i, f_i), i = 1, 2, \cdots, n,$ a third-order polynomial for each interval between knots is represented in Eq. (9).

$$s_i(x) = a_i + b_i(x - \tilde{x}_i) + c_i(x - \tilde{x}_i)^2 + d_i(x - \tilde{x}_i)^3 \tag{9}$$

A condition requires that the spline function go through the first point of the interval, yielding

$$f_i = a_i + b_i(\tilde{x}_i - \tilde{x}_i) + c_i(\tilde{x}_i - \tilde{x}_i)^2 + d_i(\tilde{x}_i - \tilde{x}_i)^3$$

which simplifies to $a_i = f_i$, meaning that coefficient in the cubic spline must match the value of the starting point of the interval. Such a fact can be reflected in Eq. (9). Given n data points, there are $n - 1$ intervals and $4(n - 1)$ unknown coefficients. $s_i(\tilde{x})$ is used to forecast resulting values of electricity consumption. Other coefficients of the polynomial can be derived by applying the condition that each of the cubic must join at the knots. For knot $i + 1$, letting $h_i = \tilde{x}_{i+1} - \tilde{x}_i$, we have Eq. (10):

$$f_i + b_ih_i + c_ih_i^2 + d_ih_i^3 = f_{i+1} \tag{10}$$

Further, the first derivatives at interior nodes must be equal, so differentiating Eq. (9) yields

$$s_i'(x) = b_i + 2c_i(x - \tilde{x}_i) + 3d_i(x - \tilde{x}_i)^2 \tag{11}$$

The equivalence of derivatives at an interior node $i + 1$ can therefore be written as

$$b_i + 2c_ih_i + 3d_ih_i^2 = b_{i+1} \tag{12}$$

The second derivatives at interior nodes must also be equal, so differentiating Eq. (11) gives

$$s_i''(x) = 2c_i + 6d_i(x - \tilde{x}_i) \tag{13}$$

The equivalence of the second derivatives at an interior node $i + 1$ can be thus expressed as

$c_i + 3d_ih_i = c_{i+1}$, namely $d_i = \dfrac{c_{i+1} - c_i}{3h_i}$. Substituting the expression into Eqs. (10) and (12) leads to

$$f_i + b_ih_i + \frac{h_i^2}{3}(2c_i + c_{i+1}) = f_{i+1} \tag{14}$$

$$b_i + h_i(c_i + c_{i+1}) = b_{i+1} \tag{15}$$

Eq. (14) can be solved for

$$b_i = \frac{f_{i+1} - f_i}{h_i} - \frac{h_i}{3}(2c_i + c_{i+1}) \tag{16}$$

The index of Eq. (16) can be reduced by 1, so

$$b_{i-1} = \frac{f_i - f_{i-1}}{h_{i-1}} - \frac{h_{i-1}}{3}(2c_{i-1} + c_i) \tag{17}$$

Eq. (15) can also be reduced by 1, *i.e.*

$$b_{i-1} + h_{i-1}(c_{i-1} + c_i) = b_i \tag{18}$$

Eqs. (16) and (17) are substituted into Eq. (18) to obtain

$$h_{i-1}c_{i-1} + 2c_i(h_{i-1} - h_i) + h_i c_{i+1} = 3\frac{f_{i+1} - f_i}{h_i} - 3\frac{f_i - f_{i-1}}{h_{i-1}}$$

Now letting, $f[\tilde{x}_i, \tilde{x}_j] = \dfrac{f_i - f_j}{\tilde{x}_i - \tilde{x}_j}$ the above equation can be rewritten as

$$h_{i-1}c_{i-1} + 2c_i(h_{i-1} - h_i) + h_i c_{i+1} = 3(f[\tilde{x}_{i+1}, \tilde{x}_i] - f[\tilde{x}_i, \tilde{x}_{i-1}]) \quad (19)$$

Moreover, considering the natural spline, the second derivative at the end of knots equals zero, which implies $s_i''(\tilde{x}_i) = s_i''(\tilde{x}_n) = 0$ Eq. (13) indicates $c_1 = 0$ and $c_{n-1} + 3d_{n-1}h_{n-1} = c_n = 0$ Consequently, we can express Eq. (19) as a matrix form by using Eq. (20). Once the vector $\{c\}$ is known, each spline can be determined accordingly.

$$\begin{bmatrix} 1 & & & & \\ h_1 & 2(h_1 + h_2) & h_2 & & \\ \ddots & \ddots & \ddots & & \\ & h_{n-2} & 2(h_{n-2} + h_{n-1}) & h_{n-1} & \\ & & & 1 & \end{bmatrix} \begin{Bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{n-1} \\ c_n \end{Bmatrix} = \begin{Bmatrix} 0 \\ 3(f[\tilde{x}_3, \tilde{x}_2] - f[\tilde{x}_2, \tilde{x}_1]) \\ \vdots \\ 3(f[\tilde{x}_n, \tilde{x}_{n-1}] - f[\tilde{x}_{n-1}, \tilde{x}_{n-2}]) \\ 0 \end{Bmatrix} \quad (20)$$

2.5 Selection of σ

To verify the performance of the model, the mean absolute percentage error (MAPE) is taken into account and defined in Eq. (21).

$$MAPE = \frac{1}{m}\left(\sum_{i=1}^{m}\left|\frac{A_i - F_i}{A_i}\right|\right) \times 100\% \quad (21)$$

where m is the number of observations, A_i is the actual value, and F_i is the forecast value.

Since σ is essential to GRNN, Eq. (21) suggests viewing MAPE as a function parameterized by σ More precisely, MAPE in the context of the same training data can be formulated as a function $f(\sigma)$ in our architecture whereby to find the critical point $\sigma \in [0,1]$ for minimizing $f(\sigma)$. Suppose that $f(\sigma)$ is continuous on the prescribed interval, such a point σ can be resolved by finding the root of derivative $f(\sigma) = 0$ and validated by the second-derivative test for $f''(\sigma) > 0$. Alternatively, the minimum of $f(\sigma)$ as well as its critical point can be obtained through numerical methods like golden-section search and parabolic interpolation. To this end, Fig. 3 shows the flowchart of our proposed approach. To determine the error of estimation, MAPE was assessed after performing GRNN. The value of σ making the least MAPE is of interest to us. We may as well mention in passing that, when it comes to MATLAB for numerical analysis, fminbnd is a readily-accessible command of avail to find the minimum of $f(\sigma)$.

As shall be clarified shortly in the following section, our proposed scheme mainly depends on the statistics of nearly 20 years (1995 ~ 2014) for training and then forecasts the consumption values of electricity of the following years, based on a sequence of recent historical data. The forecasting process is conducted over the span of years from 2015 until 2022. Also, the five-year period of forecasting from 2015 to 2019 is intended for the evaluation of the quality of the trained models by MAPE. The forecasting for the remaining years from 2020 to 2022 is regarded as the major period concerned. The value of used in the forecast for each year will be diverse because the data model of each year is diverse as well.

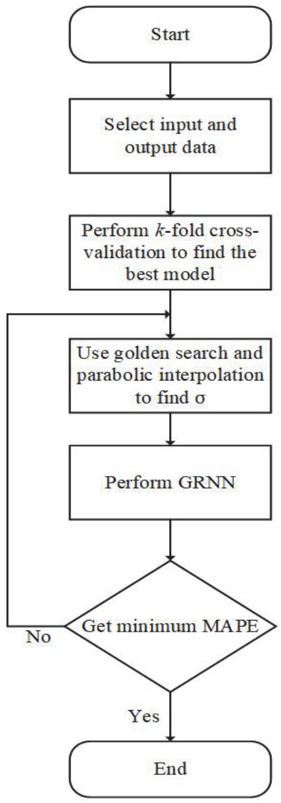

Fig. 3 Flowchart of the proposed approach

3. RESULTS AND DISCUSSION

The dataset under discussion was directly taken from the real data publicized by PLN [1] (a state-own company of Indonesia) in GWh (Gigawatt hours) unit. As summarized in Table 1, the dataset is mainly comprised of two parts: (1) load data obtained from 1995 until 2014 and (2) other data obtained from 2015 until 2019. In this study, the former is utilized for training, while in contrast, the latter is utilized for testing so as to further verify the estimation error of the proposed method. Also, two additional techniques, polynomial regression and cubic spline interpolation, are introduced to be compared. Our consistency. This was accomplished by mapping minimum and maximum values to as expressed in Eq. (22), so that the error of estimation could be minimized.

[1] PLN (PT Perusahaan Listrik Negara) is an Indonesian government-owned corporation which generates the majority of the country's electrical power and monopolizes electricity distribution nationwide.

Table 1 Electricity statistics in Indonesia

Year	Consumed (GWh)	Produced (GWh)
1995	49629	54597
1996	57000	64200
1997	64724	77065
1998	65357	74924
1999	71734	84611
2000	79170	92821
2001	84499	101630
2002	87088	108360
2003	90440	113020
2004	99826	119105
2005	105933	124505
2006	112609	131710
2007	121247	139711
2008	129018	148058
2009	134582	157337
2010	147300	169786
2011	157993	183421
2012	174342	201735
2013	188342	222207
2014	199028	238019
2015	204280	239750
2016	217438	251606
2017	226014	254660
2018	239012	267085
2019	245520	278941

Table 2 Forecast results from different GRNNs

Year	Real data	GRNN$^+$(4)	GRNN$^+$(5)	GRNN$^+$(6)
2015	204280	209805	201197	217311
2016	217438	214498	216572	233636
2017	226014	217373	228219	244470
2018	239012	222278	229467	245909
2019	245520	225269	246682	264105
MAPE	–	4.63%	1.47%	6.49%

Table 4 Forecast results for 2020 ~ 2022

Method	2020	2021	2022
GRNN$^+$(5)	246381	268704	275009
Cubic spline	246950	389483	462990
Polynomial regression	264222	295543	311864

3.1 Preparation

All data to be fed into GRNN first was subjected to a normalization process for heightened consistency. This was accomplished by mapping minimum and maximum values to as expressed in Eq. (22), so that the error of estimation could be minimized.

$$x = \frac{(y_{max} - y_{min})(x - x_{min})}{x_{max} - x_{min}} + y_{min} \tag{22}$$

where $y_{max} = 1$ and $y_{min} = -1$.

Table 3 Comparison of Subject Schemes

Year	Real data	GRNN$^+$(5)	Error (%)	Cubic spline	Error (%)	Polynomial regression	Error (%)
2015	204280	201197	1.51%	208881	2.25	208668	2.15
2016	217438	216572	0.4%	220384	1.36	221831	2.02
2017	226014	228219	0.98%	236017	4.43	235521	4.21
2018	239012	229467	3.99%	258262	8.05	249737	4.49
2019	245520	246682	1.47%	289600	17.95	264479	7.72

To forecast the values of future time steps of a sequence, we resorted to a sequence-to-sequence regression analysis, where the responses are the training sequences with values shifted by one-time step. That is, at each time step of the input sequence, our GRNN learns to forecast the value of the next time step. Now that the sequence of most recent historical data of different lengths is a determinant of regression, in our architecture the value to be forecast for the current year r is carried out with reference to data sequence (x_{r-1}, x_{r-2}, …, x_{r-t},) of past t years, t = 4,5,6. For the purpose of better presentation, in what follows, our model designating the prescribed three combinations are referred to as GRNN$^+$(4), GRNN$^+$(5), and GRNN$^+$(6), respectively. We shall investigate which of the combination could best serve our purposes.

Table 2 indicates that GRNN$^+$(5) allows the recent five years of data for better performances than those using fewer or more input data, where GRNN$^+$(5) has a MAPE of 1.47% in contrast to GRNN$^+$(4) having 4.63% and GRNN$^+$(6) having 6.49%. Figure. 4 also depicts overall trends of GRNN-forecasting results. It means that GRNN$^+$(5) forecast tends to agree with the real statistics, whereas the other two configurations appear less matched to the real data. Hereinafter, GRNN$^+$(5) shall be used as our primary means to forecast electricity consumption because it shows not only the best accuracy but also the best stability for each point.

3.2 Comparison of Subject Methods

To justify our treatment, GRNN$^+$(5) was compared quantitatively with two counterpart schemes, cubic splines and polynomial regression. Table 3 compares performance results from the subject schemes, where GRNN$^+$(5) is shown to outperform other methods in sense of maintaining comparatively lower yet more stable errors of estimation. To be more specific, GRNN$^+$(5) produces an error of estimation within 4% with respect to real values during the test phase over the five-year period from 2015 to 2019. In comparison, estimates of cubic spline and polynomial regression incur instability in that cubic splines and polynomial regression appear to have residual errors growing over time and exceeding 17% and 7%, respectively, in 2019. In addition,

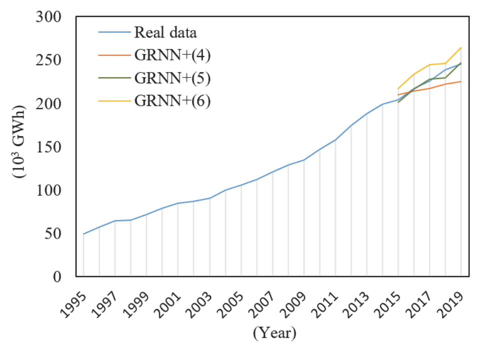

Fig. 4 GRNN-forecasting results using historical data of different lengths

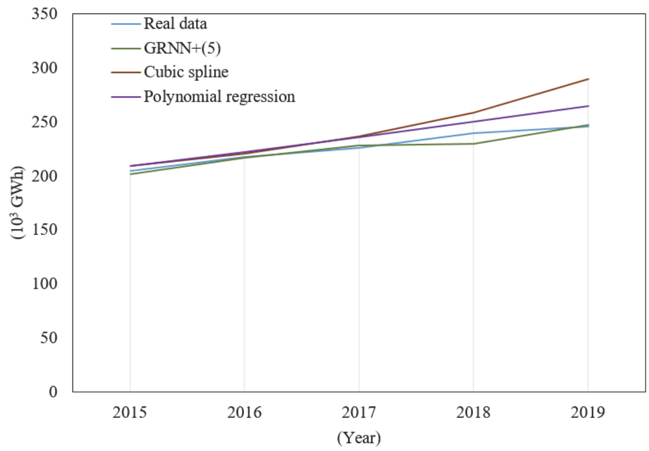

Fig. 5 A detailed view of forecast results from subject schemes during 2015 to 1019

GRNN$^+$(5), cubic splines, and polynomial regression bring about a MAPE of 1.47%, 6.81%, and 4.12%, respectively, thus implying the potency of our approach. A more detailed view of forecast results from subject schemes during the five-year period of 2015 to 2019 is illustrated in Fig. 5. The figure suggests that our approach has an inclination of matching up with real data.

After investigating the best method during the test phase, we proceeded to forecast annual electricity consumption for 2020 to 2022 by GRNN$^+$(5), which operates in synergy with *k*-fold cross-validation, golden-section search and parabolic interpolation, as outlined in Fig. 3. As mentioned earlier, *k*-fold cross-validation lends itself to less biased nor less optimistic estimates of electricity consumption, thanks to proper selections of training data. Golden-section search and parabolic interpolation were employed to find a parameter that could minimize residual errors. In this light, Table 4 lists the forecast results from our GRNN$^+$(5) relative to the other two schemes. Numerical results indicate that annual electricity consumption in Indonesia for the next three years tended to increase steadily. The projected increase in the context of cubic splines and polynomial regression appears more pronounced than that in our architecture. Discrepancies in estimates of different schemes appear more significant in 2021 and 2022. Given that our GRNN$^+$(5), compared with counterpart methods, could foresee energy consumption rises in a moderate stable manner, this is more likely reasonable to serve a capital reference base for mitigating energy supply-demand mismatch in the future.

4. CONCLUSIONS

In this study, we enhanced a GRNN model by allowing for *k*-fold cross-validation, golden-section search, and parabolic interpolation to forecast Indonesia's annual electricity consumption. The forecast was conducted on the basis of historical data of different year spans. *K*-fold cross-validation was employed to divide the dataset into 10 subsets, while, in contrast, golden-section search and parabolic interpolation helped locate the spread parameter values for optimizing the entire process. To forecast the consumption value in a certain year, a sequence of electricity data of the past five years was taken as input to yield results with a smaller error of estimation. The proposed GRNN$^+$(5) was compared with polynomial regression and cubic splines to validate its effectiveness. Quantitative comparison suggests that GRNN$^+$(5) has a good level of accuracy and stability, maintaining a low MAPE of 1.47% and thereby showing its salient strengths over counterpart schemes. This study is of value for the government or management authorities to consider and plan strategic policies of energy supply to an area in the future.

REFERENCES

Arlot, S. and Celisse, A. (2010). "A Survey of Cross-Validation Procedures for Model Selection." *Statist. Surv.*, **4**, 40-79.

Barrow, D.K., and Crone, S.F. (2016). "Cross-Validation Aggregation for Combining Autoregressive Neural Network Forecasts." *International Journal of Forecasting*, **32**(4), 1120-1137.

Cross, A.J., Rohrer, G.A., Brown-Brandl, T.M., Cassady, J.P., and Keel, B.N. (2018). "Feed-forward and Generalised Regression Neural Networks in Modelling Feeding Behaviour of Pigs in the Grow-Finish Phase." *Biosystems Engineering*, **173**, 124-133.

Fan, G.F., Wei, X., Li, Y.-T., and Hong, W.-C. (2020). "Forecasting Electricity Consumption Using a Novel Hybrid Model." *Sustainable Cities and Society*, **61**.

Filik, Ü.B., Gerek, Ö.N., and Kurban, M. (2011). "A Novel Modeling Approach for Hourly Forecasting of Long-Term Electric Energy Demand." *Energy Conversion and Management*, **52**(1), 199-211.

Heydari, A., Garcia, D.A., Keynia, F., Bisegna, F., and Santoli, L.De. (2019). "Renewable Energies Generation and Carbon Dioxide Emission Forecasting in Microgrids and National Grids Using GRNN-GWO Methodology." *Energy Procedia*, **159**, 154-159.

Jung, K., Bae, D.H., Um, M.J., Kim, S., Jeon, S., and Park, D. (2020). "Evaluation of Nitrate Load Estimations Using Neural Networks and Canonical Correlation Analysis with *K*-Fold Cross-Validation." *Sustainability (Switzerland)*, **12**(1).

Polat, Ö. and Yildirim, T. (2008). "Genetic optimization of GRNN for pattern recognition without feature extraction." *Expert Systems with Applications*, **34**(4).

Renk, T., Iankov, D., and Jondral, F.K. (2009). "Adaptive resource allocation in wireless relay networks." *IEEE Vehicular Technology Conference*, **3**, 1-5.

Rohani, A., Taki, M., and Abdollahpour, M. (2018). "A Novel Soft Computing Model (Gaussian Process Regression with *K*-Fold Cross Validation) for Daily and Monthly Solar Radiation Forecasting (Part: I)." *Renewable Energy*, **115**, 411-422.

Ruiming, F., and Shijie, S. (2020). "Daily Reference Evapotranspiration Prediction of Tieguanyin Tea Plants Based on Mathematical

Morphology Clustering and Improved Generalized Regression Neural Network." *Agricultural Water Management*, **236**.

Specht, D. F. (1991). "A General Regression Neural Network." *IEEE Transactions on Neural Networks*, **2**(6), 568-576.

Sun, Y., Lang, M., Wang, D., and Liu, L. (2014). "A PSO-GRNN model for railway freight volume prediction: Empirical study from China." *Journal of Industrial Engineering and Management*, **7**(2), 413-433.

Wan, X., Li, Y., Xia, C., Wu, M., Liang, J., and Wang, N. (2016). "A T-Wave Alternans Assessment Method Based on Least Squares Curve Fitting Technique." *Measurement*, **86**, 93-100.

Wang, Z.X., Li, Q., and Pei, L.L. (2018). "A seasonal GM(1,1) Model for Forecasting the Electricity Consumption of the Primary Economic Sectors." *Energy*, **154**, 522-534.

Wu, L., Gao, X., Xiao, Y., Yang, Y., and Chen, X. (2018). "Using a Novel Multi-Variable Grey Model to Forecast the Electricity Consumption of Shandong Province in China." *Energy*, **157**, 327-335.

Yip, H., Fan, H., and Chiang, Y. (2014). "Predicting The Maintenance Cost of Construction Equipment: Comparison Between General Regression Neural Network and Box–Jenkins Time Series Models." *Automation in Construction*, **38**, 30-38.

Zhang, Y., Niu, J., and Na, S. (2019). "A Novel Nonlinear Function Fitting Model Based on FOA and GRNN." *Mathematical Problems in Engineering*, 2019.

Zhou, Q., Jiang, H., Wang, J., and Zhou, J. (2014). "A Hybrid Model for $PM_{2.5}$ Forecasting Based on Ensemble Empirical Mode Decomposition and a General Regression Neural Network." *Science of The Total Environment*, **496**, 264-274.