Determining Optimal Operating Reserves toward Wind Power Penetration in Indonesia based on Hybrid Artificial Intelligence

DHANY HARMEIDY BARUSI, (Student Member, IEEE), AND RINALDY DALIMI1
1 Departemen Teknik Elektro, Universitas Indonesia, Depok 16424, Indonesia
Corresponding author: Rinaldy Dalimi (rinaldy@eng.ui.ac.id)

ABSTRACT The stability and economic level of the power system operation during the penetration of Wind Power Plants (WPPs) are much determined by the variability and uncertainty of the wind power output. The characteristics of seasonal wind power output can be used to define the optimal operating reserves of a stable and cost-effective power system operation. This paper proposes a comprehensive algorithm of hybrid Artificial Intelligence (AI) approach that combines the Seasonal Autoregressive Integrated Moving Average (SARIMA) and selected Neural Network Variants (NNVs) in Seasonal Daily Variability and Uncertainty (SDVU) scheme. Among all NNVs, Long Short-Term Memory (LSTM) shows the most consistent and accurate results. With the hybrid AI approach, this algorithm calculates the Dynamic Confidence Level (DCL) to determine hourly operating reserves on a daily basis. The proposed algorithm has been successfully tested using historical data of real-world WPPs that operated in Indonesia. Furthermore, the comparison toward non-seasonal with a Static Confidence Level (SCL) in several percentile scenarios is made to prove the cost-effectiveness advantages of this new algorithm that may save up to 4.2% of total daily energy consumption. An interface application is added so that the results of this research can be directly utilized by users both on the observed power system and generally in Indonesia.

INDEX TERMS dynamic confidence level, neural network, operating reserve, wind power, SARIMA.

I. INTRODUCTION

Wind Power Plant (WPP) is part of renewable energy sources, which has green energy and rapid expansion throughout the world. But integrating WPPs into a power system means taking into account all possible fluctuations caused by variability and uncertainty of wind power output. It needs additional operating reserves to maintain the stability of the power system due to the penetration of WPPs. In [1], the authors explained the reserve procurement approach and transmission capacity allocation in a power system with high penetration of WPPs. Other literature [2] describes the multi-purpose dynamic economic delivery model with renewable liability requirements. It proposed model incorporates various renewable energy resources including WPPs into the grid.

Various methodologies have been presented to calculate the number of operating reserves due to the penetration of WPPs. Based on the principle of the methods, they are classified into statistical and probabilistic, Artificial Intelligence (AI), and hybrid AI. The traditional ones typically use only statistical and probabilistic approaches to analyze the observed data. In [3], the authors used a statistical approach to determine reserve requirements in Nordic countries based on hourly wind power variations. Other studies use a combination of statistical and probabilistic approaches for that purpose [4]-[6]. AI methods include the utilization of machine learning such as neural networks and fuzzy approaches. In [7], the authors used a fuzzy power flow tool based on a steady-state security assessment to determine reserve settings due to wind power uncertainty forecast. The hybrid approach incorporates various models to get their respective benefits and deliver excellent results. In [8], the authors offer an improved Grey BP Neural Network with an Autoregressive Integrated Moving Average (ARIMA) in modeling the uncertainties of wind-wave forecast and swell forecast.

In terms of the flexibility of the reserves, [9] uses a probabilistic approach to obtain dynamic reserve scheduling.
Dynamic scheduling of operating reserves can also be used in co-optimized electricity markets with high WPPs penetration [10]. In recent years, studies on operating reserves are not only at the system level but also at the plant level. They focus only on balancing the intermittent output of the WPPs and reducing the size of reserve requirements. Most of the studies focus on the optimal size of the Energy Storage System (ESS) [11]-[13].

This paper proposes a Seasonal Daily Variability and Uncertainty (SDVU) scheme based on a hybrid AI approach that incorporates Seasonal Autoregressive Integrated Moving Average (SARIMA) and Neural Network Variants (NNVs) to obtain optimal additional operating reserves due to penetration of WPPs in the southern Sulawesi power system, Indonesia. Previously, other literature uses a similar hybrid AI approach based on a combination of SARIMA and neural network (multi-layer perceptron) for wind speed forecasting [14], however, it is not directly applicable to our study. Furthermore, it only used one type of neural network as the comparison, while in our proposed algorithm we used multiple ones.

In the proposed model, the optimal operating reserves are determined by calculating the Dynamic Confidence Level (DCL) with Hybrid AI results as the input data. Previously, [15] use probabilistic–dynamic approaches to calculate DCL. They treat expected energy not supplied as the key parameter. In this paper, we use SDVU for that purpose. We analyze that the variability and uncertainty of the observed data have a specific seasonal nature, so they can be approached linearly by SARIMA, then NNV is used to catch the non-linear nature of the SARIMA output to increase the accuracy of the final result. This study uses commonly NNV methods for time series analysis and forecast such as a Non-linear Autoregressive Neural Network (NARNET), Multi-Layer Perceptron Back Propagation (MLPBP), Wavelet Neural Network (WNN), and Long Short-Term Memory (LSTM). The NNV methods are combined one by one with SARIMA to generate forecasted models of the variability and uncertainty from the observed data. The best combination will be chosen as the proposed method in this study. It is determined based on the results of accuracy analysis using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) parameters. To prove the cost-effectiveness of this algorithm, the results of additional operating reserves are compared with the ones from Static Confidence Level (SCL) for several percentile scenarios. Table I shows the comparison of the research works on some reference papers cited previously. It notifies the correlation to the previous ones and justifies the novelty of this research.

| Authors         | Proposed Methods | Objectives       |
|-----------------|------------------|------------------|
| Wu et al [8]    | Grey BP NN + ARIMA | wind-wave forecast |
| Wang et al [11] | Probabilistic Distribution | energy storage systems |
| Alencar et al [14] | SARIMA + MLPBP | wind speed forecast |
| Rahman et al [15] | Probabilistic-Dynamic | WPPs dynamic reserves |
| Barus & Dalimi  | SARIMA + LSTM | WPPs DCL reserves |

The question to be answered from the research is as follows: How to maintain the level of system reliability with a more economical cost during WPPs penetration?

The main contributions of this manuscript can be summarized as follows:
1) Presenting a hybrid forecasting method based on SARIMA and NNV for specific variability and uncertainty characteristics;
2) Developing an effective forecasting approach to calculate DCL;
3) Applying SDVU scheme for more economical operating reserves but still maintaining the same level of system reliability;
4) Use the real-world Indonesia database for operating reserves forecasting due to WPPs penetration, which has great potential in its future development.

The rest of this paper is organized as follows: Section II describes the observed data used in this study. The proposed methodology is presented in Section III. In Section IV, the result of SARIMA, NNV, and DCL are explored. Finally, the conclusion is provided in Section V.
II. DATASET OF OBSERVED POWER SYSTEM

This section displays the dataset of the observed power system that used in the research. It consists of system overview, historical seasonal data, and historical intermittent data derived from the statistical data of the southern Sulawesi power system.

A. SYSTEM OVERVIEW

Indonesia lies in the tropical climate area between latitudes 11°S and 6°N, and longitudes 95°E and 141°E. Fig. 1(a) shows the location of Indonesia in the southeast Asia region based on a wind potential map [16]. Sulawesi is one of the major islands in Indonesia. It has 2 separate major interconnected power systems, which are southern Sulawesi and northern Sulawesi power systems. As shown in Fig. 1(b), there are 2 WPPs connected to the southern Sulawesi power system, which are Sidrap WPP (70 MW) and Jeneponto WPP (60 MW). Sidrap WPP is located at a higher altitude with hilly terrain, while the Jeneponto WPP is at a lower one with flat coastline terrain. In [17], the authors have described the detailed technical specification of both WPPs. Table II shows the composition of power plants (PPs) in the southern Sulawesi power system. The highest peak load is 1,411 MW (Nov 2019). It means there are 379 MW (27%) reserve margins of the system, mostly in gas PP and diesel PP.

| Type of Power Plant | Capacity (MW) | Portion (%) | Role of Power Plants |
|---------------------|---------------|-------------|---------------------|
| Coal Steam PP       | 940           | 53          | Base Load           |
| Hydro PP            | 260           | 15          | Peaker              |
| Wind PP             | 130           | 7           | Must Run            |
| Gas PP              | 363           | 20          | Load Follower       |
| Diesel PP           | 97            | 5           | Standby             |
| TOTAL               | 1,790         | 100         |                     |

B. HISTORICAL SEASONAL DATA

The historical database [18] used in this study is obtained from the southern Sulawesi load dispatch center. This study uses the aggregation power output data of the two WPPs (southern Sulawesi WPPs) on an hour-ahead forecast basis. The data used has been normalized for aggregation purposes. The time-series data is from April 01, 2019, to March 31, 2020, with a time resolution of 30 minutes. The timing is to refer to the prevailing season period in Indonesia. Annually, BMKG (Badan Meteorologi Klimatologi dan Geofisika), as the National Meteorological, Climatological, and Geophysical Agency in Indonesia, publishes online documents regarding two seasonal climate forecasts, namely the dry season forecast in March [19] and the wet season forecast in September [20]. Generally, in Indonesia, the dry season period occurs between April-September and the wet season between October-March. Fig. 2 shows the annual characteristic of southern Sulawesi WPP during the seasonal period from April 01, 2019, to March 31, 2020. It uses the following parameters: Capacity Factor (CF), Daily Capacity Load Factor (DCLF), and Standard Deviation Error (SDE) to show the seasonal characteristic of the southern Sulawesi WPPs. The formulation of these parameters is as follows [21]. Within a certain period, CF is defined as the ratio of the total actual energy produced ($E_{act}$) to the maximum energy that might be generated if the power plant operates at full load based on its installed capacity ($E_{max}$). By using (1), CF may also be calculated as the average power generated ($\bar{P}$) divided by the installed capacity of the power plant ($P_{inst}$).

$$CF = \frac{E_{act}}{E_{max}} = \frac{\bar{P}}{P_{inst}} \quad (1)$$

For daily operating patterns of WPP in a certain period, the DCLF is another useful indicator to note. As defined in (2), $DCLF$ is the average CF ($\bar{CF}$) divided by the maximum CF ($CF_{max}$) in a day (24 hours) for a specified period.

$$DCLF = \frac{\bar{CF}}{CF_{max}} \quad (2)$$
As the basis [22], the forecast deviation ($P_{dev}$) means the difference between the actual ($P_{act}$) and the forecasted value ($P_{fc}$) with time function ($t$), which defined as:

$$P_{dev}(t) = P_{act}(t) - P_{fc}(t)$$

(3)

Standard Deviation Error ($SDE$) which refers to the standard deviation of the error distribution, is defined as the average root square error deviation for the entire observed duration ($n$) and calculated for each time function (4).

$$SDE(t) = \left( \frac{1}{n-1} \sum_{t=1}^{n} (P_{dev}(t) - P_{dev}(t))^2 \right)^{1/2}$$

(4)

Fig. 2(a) shows that $CF$ in the dry season period is higher than in the rainy season, with the highest $CF$ occurring in July 2019 and the lowest in December 2019. It shows that southern Sulawesi WPPs are more productive during the dry season. Accordingly, $DCLF$ in the dry season is higher than in the rainy season. This shows that the WPPs energy production during the dry season is more stable than the rainy one.

In terms of forecast accuracy, $SDE$ in the rainy season is almost the same as the one in the dry season. It shows that the accuracy of forecasting is almost evenly distributed throughout the year and does not depend on the season period.

Fig. 2(b) shows that, on a daily basis, the highest $CF$ occurs at around 3 pm, either on an annual period, dry season, or rainy season. It also presents that $DCLF$ in the dry season is much higher than in the rainy season. It confirms the previously displayed data.

From these two conditions, it can be seen that there are seasonal characteristics of the WPPs output data, both in terms of production patterns and prediction accuracy. These results will be used as the basis for using the SARIMA method in a more comprehensive analysis of the linear properties of the observed data.

C. HISTORICAL INTERMITTENT DATA

WPP is part of Variable Renewable Energy (VRE) with specific intermittent characteristics. It has natural changes in the level of availability over time (variability) that may not be perfectly forecasted (uncertainty) [23]. The variability of wind power is due to the natural fluctuation of wind power output which is an effect of the availability of natural resources. It still exists even if the forecast is accurate. WPP uncertainty can be defined as the deviation that occurs between forecasted and actual wind power output. The accuracy of the WPP forecast can be determined by some elements such as forecast horizons and geographic distribution. As WPP contribution increases, additional operating reserves are required due to the increase of variability and uncertainty that appeared in the power system [24]-[25].

In this study, wind variability means the difference between 30 minutes WPP output series, while wind uncertainty means the difference between the 30 minutes of the actual WPP output and the estimated value in the time frame of 1 hour ahead. Fig. 3 describes wind variability and uncertainty.

![FIGURE 3. Historical Wind Variability and Uncertainty from Southern Sulawesi WPP, Indonesia](image-url)
historical data of south Sulawesi WPP in the observed period. It shows the white noise characteristic of the data that describes the high complexity and stochastic level. This characteristic will be observed using the proposed method that will be discussed more in the next chapters. SARIMA will handle the linear nature of the data, while several NNVs will be tested to explore their non-linearity and weighting their accuracy and consistency based on commonly used statistical test parameters such as MAE, RMSE, and MAPE. The database used included 366 sampling days every 30 minutes, totaling 17,568 points for each variable and a total of 35,136 points.

III. PROPOSED METHODOLOGY
The Seasonal Daily Variability and Uncertainty (SDVU) scheme is the proposed methodology in this study. The role of SDVU in the daily production simulation procedure can be seen in Fig. 4(a). SDVU takes input data from the daily wind power forecast phase which then its results are used as input for the economic dispatch simulation phase. In this study, fluctuations due to load forecast and uncertainties due to the Forced Outage Rate (FOR) of generation and transmission systems are assumed to have been calculated in the hydrothermal coordination phase. The system level is chosen in this study to calculate the number of additional operating reserves due to penetration of WPPs by considering system stiffness.

Details of the SDVU process scheme can be seen in Fig. 4(b). First, a large database is created from seasonal wind power behavior as initial input [18]. Variability and uncertainty variables are extracted from the large database that represented WPP aggregation. Then, the seasonal characteristic of data is approached linearly using the SARIMA method. Within the SARIMA process, clustering and Principal Component Analysis (PCA) steps are included to minimize the white noise effect of existing data. After that, NNV will catch the non-linear relations of SARIMA results to produce a more accurate forecast. The best combination of SARIMA and NNV will be used as the proposed Hybrid AI method for this study. It is determined based on the results of the accuracy analysis with MAE, RSME, and MAPE parameters. Furthermore, with a statistical approach, the DCL quantities will be determined based on the results of the proposed method. Then, it will be used to calculate the hourly additional operating reserve for several percentile scenarios on a daily basis. To prove the cost-effectiveness advantage of the proposed method, comparisons are made to the non-seasonal method based on the Static Confidence Level (SCL), regarding the need for additional operating reserves and associated energy saving for several percentile scenarios.

A. OVERVIEW OF SARIMA METHOD
The SARIMA method uses a statistical approach to predict value as a linear function of the observed data. The method is a derivative of the ARIMA model which is usually utilized to examine data with a time-series pattern that has seasonal characteristics [26]. From the previous analysis [18], it was known that the wind variability and uncertainty of the observed data had a specific seasonal characteristic, so SARIMA is suitable to be used in this study.

The SARIMA model is usually displayed as \( \text{SARIMA} (p,d,q)(P,D,Q)_s \). The non-seasonal parts of the existing model are denoted by lowercase notations, while uppercase ones denote the seasonal part of the observed model, where \( s \) denotes the number of periods in a season. For instance, the weekly sequence of days is displayed as \( s=7 \).
Notation $p$ and $P$ denote the autoregressive parameter, $d$ and $D$ denote the difference parameter, and $q$ and $Q$ denote the moving average parameter.

The mathematical formulation of the SARIMA Model is as follows:

$$
\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^S)^Dy_t = \theta_q(B)\Theta_Q(B^S)\epsilon_t
$$

where $\phi$ is an AR parameter of order $p$, $\theta$ is a MA parameter of order $q$, $\Phi$ is a seasonal AR parameter of order $P$, $\Theta$ is a seasonal MA parameter of order $Q$, $(1-B)^d$ is the differentiator parameter, $(1-B^S)^D$ is the seasonal differenitator parameter, $B$ is the backshift parameter as $B^k(y_t) = y_{t-k}$, $\epsilon_t$ is the residual deviation at time $t$, and $y_t$ is the variability or uncertainty for $t$.

$$
\phi_p(B^S) = 1 - \Phi_1(B^S) - \Phi_2(B^{2S}) - \ldots - \Phi_p(B^{ps})
$$

$$
\theta_Q(B^S) = 1 - \Theta_1(B^S) - \Theta_2(B^{2S}) - \ldots - \Theta_Q(B^{qs})
$$

$$
\phi_p(B) = 1 - \phi_1(B) - \phi_2(B^2) - \ldots - \phi_p(B^p)
$$

$$
\theta_q(B) = 1 - \theta_1(B) - \theta_2(B^2) - \ldots - \theta_q(B^q)
$$

Equations (5)-(9) will be used to analyze the linearity portion of the seasonal observation data. This is the first stage in the SDVU method.

The main stages in obtaining the SARIMA model consist of four steps: identify, estimate, diagnose, and forecast [26]. But in this study, we include clustering, PCA, and regroup as additional steps. The complete steps are as follows:

1) Clustering time-series data into 24 groups based on data clusters per 1 hour.
2) Use the principal component analysis (PCA) method [27] on the clustering results as input data for the SARIMA method.
3) Identify the existence of seasonality data with the Auto Correlation Functions (ACF) and Partial Auto Correlation Functions (PACF).
4) Estimated coefficients of the models created.
5) Diagnose model validity with standard accuracy parameters (R-square, RMSE, MAPE, MAE, and normalized Bayesian information criterion).
6) Perform a forecast based on the predetermined model
7) Regroup the time-series result which is then be used as input for the NNV methods.

B. OVERVIEW OF NNV METHOD

The Neural Network (NN) method is designed using performance characteristics similar to the human nerve impulse delivery process [28]. This method is increasingly developing along with its ability to solve many problems that occur in various fields, such as voice recognition, face identification, and weather predictions. This study uses the NNV methods which are commonly used for time series analysis and forecast, namely: the Non-linear Autoregressive Neural Network (NARNET) [28], Multi-Layer Perceptron Back Propagation (MLPBP) [14], Wavelet Neural Network (WNN) [29], and Long Short-Term Memory (LSTM) [30].

In general, the network architecture used in NNV consists of an input layer, hidden layer, and output layer. Each layer is made up of units called neurons. These relationships and their weights and biases are formulated as shown in (10) and (11) [31].

$$
v_k = \sum_{j=1}^{m} w_{kj}a_j + b_k
$$

$$
z_k = \varphi(v_k)
$$

Where, $w_{kj}$ is the weight moving from input $k$ to the hidden neuron $j$, $z_k$ is the neuron output, $a_j$ is the neuron input, $b_k$ is the bias, and $\varphi$ is the activation function.

At the training phase, an iterative learning process occurs. The network processes the existing input and output data by adjusting weights and biases on an ongoing basis to obtain mismatches that are smaller than the allowable tolerances or the number of iterations reaching the maximum epoch. Model validation is carried out after training is complete to ensure the process produces accurate output.

The complete steps in this phase are as follows:

1) Preparation of SARIMA output data as input for the selected NNV methods (NARNET, MLPBP, WNN, LSTM).
2) Determination of the initial configuration and options for each NNV methods (training methods, number of hidden layers, max epoch, etc.).
3) Scripting for possible scenario configurations (credible options) for each NNV.
4) Perform the training, validating, testing, and forecasting processes for each method of the NNV and record the statistical accuracy parameters (MAE, MAPE, RMSE) for each scenario.
5) Determination of the proposed hybrid AI method based on the most accurate result of selected NNV methods.

C. OVERVIEW OF ACCURACY EVALUATION METHOD

Several accuracy evaluation methods were adopted in this study [22], namely: mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE). This accuracy index calculated as follows:

$$
MAE(t) = \frac{1}{n} \sum_{t=1}^{n} |P_{dev}(t)|
$$

$$
RMSE(t) = \left( \frac{1}{n} \sum_{t=1}^{n} P_{dev}^2(t) \right)^{\frac{1}{2}}
$$
MAE, RMSE, and MAPE are statistical parameters that are commonly used in various references to calculate the accuracy level of an error prediction in time series analysis based on absolute, average, and relative aspects.

D. OVERVIEW OF DCL PARAMETER

By using a statistical approach, we use historical data of wind variability and uncertainty to determine the size of the additional operating reserves needed in the penetrated system. It commonly uses the n-sigma criterion \([15], [32]\). If wind variability and uncertainty are assumed to be uncorrelated, the wind intermittency sigma \((\sigma_{\text{int}})\) of the time series data will be:

\[
\sigma_{\text{int}} = \sqrt{\sigma_{\text{var}}^2 + \sigma_{\text{unc}}^2}
\]  

Where, \(\sigma_{\text{var}}\) is the sigma of wind variability and \(\sigma_{\text{unc}}\) is the sigma of wind uncertainty. The greater the \(\sigma_{\text{int}}\) value, the greater the additional reserves that need to be provided. In this study, the amount of this reserve associated with \(\sigma_{\text{int}}\) magnitude is defined as the Confidence Level (CL) since it shows the level of confidence of the penetrated system. In [2], [3], [29] the Static Confidence Level (SCL) is used which assumes a fixed sigma magnitude over the period. In this study, Dynamic Confidence Level (DCL) is introduced, which come from the hourly \(\sigma_{\text{int}}\) magnitudes that vary over the observed period. These dynamic quantities are obtained from forecasted variability and uncertainty variables using the Hybrid AI method.

The Probability Density Function (PDF) of the observed or forecasted data can be used to determine the sigma magnitude of its variability and uncertainty. Fig. 5 describes the pdf of a normal distribution [33]. It shows the correlation between the size of the sigma multiplier \((n)\) and the probability of occurrence (percentile). The greater the value of the sigma multiplier \((n)\), the greater the percentile.

In the context of WPP penetration, the additional operating reserve value \((\Delta Res)\), as shown in (16), is the \(\sigma_{\text{int}}\) the magnitude of the observed data times the multiplier \((n)\) of the desired percentile level.

\[
\Delta Res = n. \sigma_{\text{int}}
\]  

The larger the percentile chosen means the greater the WPP fluctuation can be reduced (the penetrated system is more stable and reliable), but consequently the more additional operating reserves that need to be prepared.

IV. RESULTS OF THE SIMULATION

This section displays the results attained from the proposed methodology using SDVU based on hybrid AI. Observed data were taken from the aggregation of the southern Sulawesi WPP for a one-year duration including the dry and wet season period in Indonesia of tropical climate area. Forecasting models were developed using SPSS modeler software for the SARIMA method [34] and MATLAB statistics and machine learning software for the NNV method [35]. The results are presented as follows.

A. RESULT OF SARIMA METHOD

This section presents the forecasting results obtained from the SARIMA method process. The first stage is clustering variability and uncertainty time-series data into 24 clusters according to the existing time scaling (1-hour interval). Then, the PCA process is carried out on the data. The aim is to reduce the white noise characteristic of existing stochastic data. SARIMA model parameters are then automatically obtained from the expert modeler feature in the SPSS software. The selection of the best-fitting model refers to ACF, PACF, as well as statistical accuracy parameters, such as R-square, MAE, RMSE, MAPE, and normalized Bayesian information criterion. The best SARIMA configuration is displayed in Table IV. It shows that the weekly cycle on a daily basis \((s=7)\) is the best seasonal period for the observed data.
### TABLE IV

| Time | SARIMA Model for Variability and Uncertainty |
|------|---------------------------------------------|
| 1    | (p, d, q) (P, D, Q)                       |
| 2    | 3                                           |
| 4    | 5                                           |
| 6    | 7                                           |
| 8    | 9                                           |
| 10   | 11                                          |
| 12   | 13                                          |
| 14   | 15                                          |
| 16   | 17                                          |
| 18   | 19                                          |
| 20   | 21                                          |
| 22   | 23                                          |
| 24   |                                             |

Var:
- (2,1,1): (1,1,1), (0,0,0), (4,1,1), (2,1,1), (1,1,1), (0,1,1)
- (0,1,0,2,1): (1,1,0), (0,1,2), (1,1,3), (1,1,1), (1,1,0)
- (2,0,1): (0,1,1), (2,0,1), (3,0,1), (2,0,3), (2,0,1), (0,1,1)
- (2,0,1,2,0): (1,1,0), (0,1,2), (3,0,1), (2,1,1), (1,2,3), (2,0,1)
- (2,0,4,1,2): (0,1,1), (1,1,1), (0,1,1), (0,0,0), (0,0,1), (0,2,0)
- (2,0,1,2,0): (0,1,1), (2,0,2), (1,0,3), (0,1,2), (1,0,1), (0,1,0)

Unc:
- (3,0,1,0,1): (2,1,1), (0,0,2), (0,1,2), (2,0,0), (2,0,1), (0,1,0)
- (2,1,0,0,1): (2,2,0), (0,0,0), (3,0,1), (5,1,0), (4,0,1), (0,1,2)
- (2,0,1,3,1,0): (4,1,2), (1,2,0), (0,0,1), (0,2,0), (3,0,1), (0,1,0)
- (2,0,2,0,2): (0,1,1), (0,0,0), (0,0,3), (0,1,4), (1,0,1), (0,2,0)
- (2,0,1,0,0): (2,1,3), (1,2,0), (0,0,1), (0,1,0), (0,2,1), (0,2,0)
- (2,1,4,1,2,5): (1,0,1), (1,2,0), (2,2,1), (0,0,0), (0,2,1), (3,0,4)

### TABLE V

| Neural Network Variant (NNV) Configuration |
|-------------------------------------------|
| NNV methods | max hidden layer | max epoch |
|---------------|-----------------|-----------|
| NARNET        | closed, trainlm | 50        |
| MLPBP         | logsig, purelin, trainlm | 60        |
| WNN           | CWT+CNN, sgd, CPU | 100       |
| LSTM          | adam, CPU       | 250       |

For all NNV models, the database is separated into two groups, 80% for training and the last 20% for testing. Fig. 6 illustrates hourly variability and uncertainty on a daily basis of observed and forecasted values of each tested hybrid AI method. Table VI and Table VII show the accuracy performance results of those methods. Both displays show that SARIMA+LSTM is the most accurate and consistent hybrid AI method tested in this study.

### TABLE VI

| Hybrid AI Methods | Accuracy Indicators |
|-------------------|---------------------|
| SARIMA            | MAE 0.11443, RSME 0.14505, MAPE 34.27% |
| SARIMA + NARNET   | MAE 0.08161, RSME 0.09743, MAPE 8.11% |
| SARIMA + MLP BP   | MAE 0.06759, RSME 0.08059, MAPE 9.72% |
| SARIMA + LSTM     | MAE 0.02403, RSME 0.03643, MAPE 2.91% |
| SARIMA + WNN      | MAE 0.07174, RSME 0.08408, MAPE 12.31% |

### TABLE VII

| Hybrid AI Methods | Accuracy Indicators |
|-------------------|---------------------|
| SARIMA            | MAE 0.21534, RSME 0.39142, MAPE 42.75% |
| SARIMA + NARNET   | MAE 0.13448, RSME 0.20683, MAPE 12.46% |
| SARIMA + MLP BP   | MAE 0.22928, RSME 0.35588, MAPE 19.27% |
| SARIMA + LSTM     | MAE 0.01559, RSME 0.03151, MAPE 1.79% |
| SARIMA + WNN      | MAE 0.19172, RSME 0.25269, MAPE 11.35% |

### FIGURE 6

Hourly Variability and Uncertainty of Observed and Forecasted Values of Each Tested Hybrid AI Methods in Daily Basis
In the next section, SARIMA+LSTM becomes the proposed hybrid AI method to calculate DCL and additional operating reserves using the SDVU scheme.

C. RESULT OF DCL CALCULATION

By using (15), (16), Table II, and forecasted results of the proposed hybrid AI method (SARIMA+LSTM), we calculate DCL in the range of percentile magnitudes. Fig. 7 shows SCL versus DCL in determining the average additional daily operating reserve per unit (p.u.) of the observed data based on percentile value ranges. SCL used the maximum $\sigma_{\text{int}}$ the magnitude from the previous day while DCL uses forecasted ones from the proposed hybrid AI method.

Furthermore, this paper evaluates the cost-effectiveness advantages of DCL by comparing its results with the SCL one. The comparison is made by using Table I, Fig. 6, and Fig. 7, and selected percentile scenarios.

Figure 8 shows the curve of wind power output on a daily basis with additional daily operating reserves of SCL versus DCL for selected percentile scenarios and Table VII shows the associated energy saving for each percentile scenario.

It can be inferred that using SDVU-based DCL can produce significant savings with the same level of reliability compared to SCL. Even the P-99 with DCL requires additional operating reserve which is only about one-fourth of the P-90 with the SCL.

### TABLE VIII

| Percentile | SCL (MWh) | DCL (MWh) | Energy Saving (MWh) | Daily Saving |
|------------|-----------|-----------|---------------------|-------------|
| P-90       | 805.39    | 127.81    | 677.58              | 2.7%        |
| P-95       | 959.62    | 152.29    | 807.33              | 3.2%        |
| P-97       | 1,062.43  | 168.60    | 893.83              | 3.5%        |
| P-99       | 1,263.17  | 200.46    | 1,062.71            | 4.2%        |
D. USER INTERFACE APPLICATION

Since the result of this research was used directly by users at the Southern Sulawesi Load Dispatch Center, we created a simple user interface application so that users can use it for daily operational purposes. By using this application, the user can choose the desired day and make changes to the peak load, load factor, composition of peakers, load-followers, base-loaders, and WPPs as needed. Based on this choice, the stiffness of the system can be calculated using a statistical approach. Based on the system-level approach, the optimal percentile, total additional energy requirements, and additional hourly operating reserve composition are calculated according to the pre-selected day data. The application is designed in such a way that it can be used by generally in Indonesia, which has a typical tropical climate similar to conditions in South Sulawesi. Figure 9 shows an example of the application display to calculate the need for additional operating reserves on Thursday which is the highest peak load on each week.

V. CONCLUSION

This paper has demonstrated the implementation of Seasonal Daily variability and Uncertainty (SDVU) schemes to determine the optimal additional operating reserve in a power system due to the penetration of WPP. The proposed hybrid AI method uses the strengths of SARIMA and LSTM schemes, which can catch linear and nonlinear system characteristics. It is the best hybrid AI among all NNV methods based on the accuracy and consistency rate.

The research question in the Introduction section has been answered clearly since this innovative method is conclusively proven to give more economical additional operating reserves with the same level of system reliability during WPP penetration. With a Dynamic Confidence Level (DCL) it may save up to 4.2% of the total energy compared to the one with a Static Confidence Level (SCL). To further complement the results of this study an interface application is added so the benefit of this new breakthrough can be used in the southern Sulawesi power system, generally in other Indonesia areas or around the world with similar tropical climate characteristics. Further work to be done is to determine system-level operating reserves in the observed system based on Hybrid AI that considers load fluctuation as well as uncertainties of generation and transmission systems.

ACKNOWLEDGMENT

The authors would like to express their gratitude and appreciation to Suroso Isnandar, Nurdin Pabi, Titi Yustiana, Goklas Batapanmuda, and Tommy Kaendo from PLN UIKL Sulawesi for providing the operational data. The authors would like to thank Teddy Surya Gunawan and Pramod Jain for enriching this article with some insights.

REFERENCES

[1] N. Viafora, S. Delikaraoglu, P. Pinson, G. Hug and J. Holboll, “Dynamic Reserve and Transmission Capacity Allocation in Wind-Dominated Power Systems,” in IEEE Transactions on Power Systems, vol. 36, no. 4, pp. 3017-3028, July 2021, doi: 10.1109/TPWRS.2020.3043225.

[2] T. G. Hlalele, R. M. Naidoo, J. Zhang and R. C. Bansal, “Dynamic Economic Dispatch With Maximal Renewable Penetration Under Renewable Obligation,” in IEEE Access, vol. 8, pp. 38794-38808, 2020, doi: 10.1109/access.2020.2975674.

[3] H. Holttinen, “Impact of hourly wind power variations on the system operation in the Nordic countries,” Wind Energy, vol. 8, no. 2, pp. 157-178, 2005, doi:10.1002/we.143.

[4] M. A. Matos and R. J. Bessa, “Setting the Operating Reserve Using Probabilistic Wind Power Forecasts,” IEEE Transactions on Power Systems, vol. 26, no. 2, pp. 594-603, May 2011, doi: 10.1109/TPWRS.2010.2065818.

[5] X. Zhu, C. Liu, C. Su and J. Liu, “Learning-Based Probabilistic Power Flow Calculation Considering the Correlation Among Multiple Wind Farms,” in IEEE Access, vol. 8, pp. 136782-136793, 2020, doi: 10.1109/access.2020.301151.

[6] X. Yang, Y. Yang, Y. Liu and Z. Deng, “A Reliability Assessment Approach for Electric Power Systems Considering Wind Power Uncertainty,” in IEEE Access, vol. 8, pp. 12467-12478, 2020, doi: 10.1109/access.2020.2966275.

[7] R. J. Bessa et al., “Reserve Setting and Ready-State Security Assessment Using Wind Power Uncertainty Forecast: A Case Study,” IEEE Transactions on Sustainable Energy, vol. 3, no. 4, pp. 827-836, Oct. 2012, doi: 10.1109/TSTE.2012.2199340.

[8] F. Wu, R. Jing, X.-P. Zhang, F. Wang and Y. Bao, “A Combined Method of Improved Grey BP Neural Network and MEEMD-ARMA for Day-Ahead Wave Energy Forecast,” in IEEE Transactions on Sustainable Energy, vol. 12, no. 4, pp. 2404-2412, Oct. 2021, doi: 10.1109/TSTE.2021.309691.

[9] K. De Vos and J. Driesen, “Dynamic operating reserve strategies for wind power integration,” IET Renewable Power Generation, vol. 8, no. 6, pp. 598-610, Aug. 2014, doi: 10.1049/iet-rpg.2013.0209.

[10] K. Das et al., “Adequacy of frequency reserves for high wind power generation,” IET Renewable Power Generation, vol. 11, no. 8, pp. 1286-1294, 28 6 2017, doi: 10.1049/iet-rpg.2016.0501.

[11] C. Wang, Z. Liang, J. Liang, Q. Teng, X. Dong, and Z. Wang, “Modeling the temporal correlation of hourly day-ahead short-term wind power forecast error for optimal sizing energy storage system,” International Journal of Electrical Power & Energy Systems, vol. 98, pp. 373–381, Jun. 2018 doi: 10.1016/j.ijepes.2017.12.012.

[12] E. Oh and H. Wang, “ Reinforcement-Learning-Based Energy Storage System Operation Strategies to Manage Wind Power Forecast Uncertainty,” IEEE Access, vol. 8, pp. 20965–20976, 2020, doi: 10.1109/access.2020.3068841.

[13] X. Yu, X. Dong, S. Pang, L. Zhou, and H. Zang, “Energy Storage Sizing Optimization and Sensitivity Analysis Based on Wind Power Forecast Error Compensation,” Energies, vol. 12, no. 24, pp. 4755, Dec. 2019, doi: 10.3390/energy.12244755.

[14] D. B. Alencar, C. M. Affonso, R. C. L. Oliveira, and J. C. R. Filho, “Hybrid Approach Combining SARIMA and Neural Networks for Multi-Step Ahead Wind Speed Forecasting in Brazil,” IEEE Access, vol. 6, pp. 55986–55994, 2018, doi: 10.1109/access.2018.2872720.

[15] C. Rahmann, A. Heinemann, and R. Torres, “Quantifying operating reserves with wind power: towards probabilistic–dynamic approaches,” IET Generation, Transmission & Distribution, vol. 10, no. 2, 2015, doi:10.1049/iet-gtd.2015.0538.

[16] N. Lee, R. Cardoso de Oliveira, B. Roberts, J. Katz, T. Brown, and F. Flores-Espino, “Exploring Renewable Energy Opportunities in Select Southeast Asian Countries: A Geospatial Analysis of the Levelized Cost of Energy of Utility-Scale Wind and Solar Photovoltaics,” USAID-NREL partnership, Jun. 2019, 4-8, doi:10.2172/1527336.

[17] D. H. Barus and R. Dalimi, “The Integration of Wind Power Plants to Southern Sulawesi Power System,” in Proc. ICEECE’20, Istanbul, Turkey, 2020, pp. 643-648, doi: 10.1109/iceece49384.2020.9179480.

[18] D. H. Barus and R. Dalimi, “Multi-Stage Statistical Approach to Wind Power Forecast Errors Evaluation: a Southern Sulawesi Case
Study,” International Journal on Advanced Science, Engineering and Information Technology, vol. 11, no. 2, pp. 633–641, Apr. 2021, doi: 10.18517/ijaseit.11.2.12385.

[19] BMKG Official Website, “Prakiraan Musim Kemarau 2019 di Indonesia” Accessed: June. 24, 2020. [Online]. Available: http://www.bmkg.go.id/

[20] BMKG Official Website, “Prakiraan Musim Hujan 2019/2020 di Indonesia” Accessed: June. 25, 2020. [Online]. Available: http://www.bmkg.go.id/

[21] L. Lledó, V. Torralba, A. Soret, J. Ramon, and F. J. Doblas-Reyes, “Seasonal forecasts of wind power generation,” Renewable Energy, vol. 143, pp. 91–100, Dec. 2019, doi: 10.1016/j.renene.2019.04.135.

[22] H. Madsen, P. Pinson, G. Kariniotakis, H. A. Nielsen, and T. S. Nielsen, “Standardizing the Performance Evaluation of Short-Term Wind Power Prediction Models,” Wind Engineering, vol. 29, no. 6, pp. 475–489, Dec. 2005, doi: 10.1260/030952405776234599.

[23] E. Ela and M. O’Malley, “Studying the Variability and Uncertainty Impacts of Variable Generation at Multiple Timescales,” IEEE Transactions on Power Systems, vol. 27, no. 3, pp. 1324–1333, Aug. 2012, doi: 10.1109/tpwrs.2012.2185816.

[24] M. Hedayati-Mehdiabadi, P. Balasubramanian, K. W. Hedman, and J. Zhang, “Market Implications of Wind Reserve Margin,” IEEE Transactions on Power Systems, vol. 33, no. 5, pp. 5161–5170, Sep. 2018, doi: 10.1109/tpwrs.2018.2817649.

[25] P. Sanguansat, Ed., “Two-dimensional principal component analysis and its extensions,” in Principal component analysis, IntechOpen, Mar. 2012, pp. 1-20, doi: 10.5772/2340.

[26] G. Zhang, H. Zhou, C. Wang, H. Xue, J. Wang and H. Wan, “Forecasting Time Series Albedo Using NARnet Based on EEMD Decomposition,” in IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 5, pp. 3544-3557, May 2020, doi: 10.1109/tgrs.2019.2958048.

[27] M. Ray, K. N. Singh, V. Ramasubramanian, R. K. Paul, A. Mukherjee, and S. Rathod, “Integration of Wavelet Transform with ANN and WNN for Time Series Forecasting: an Application to Indian Monsoon Rainfall,” National Academy Science Letters, vol. 43, no. 6, pp. 509–513, Jan. 2020, doi: 10.1007/s40009-020-00887-2.

[28] H. Iba, “Meta-heuristics, Machine Learning, and Deep Learning Methods” in Evolutionary Approach to Machine Learning and Deep Neural Networks, Singapore: Springer, 2018, pp. 27-75, doi: 10.1007/978-981-13-0200-8.

[29] H. Holttinen, M. Milligan, B. Kirby, T. Acker, V. Neimane, and T. Molinski, “Using Standard Deviation as a Measure of Increased Operational Reserve Requirement for Wind Power,” Wind Engineering, vol. 32, no. 4, pp. 355–377, Jun. 2008, doi: 10.1260/0309-524x.32.4.355.

[30] Y. Li, Z. Zhu, D. Kong, H. Han, and Y. Zhao, “EA-LSTM: Evolutionary attention-based LSTM for time series prediction,” Knowledge-Based Systems, vol. 181, p. 104785, Oct. 2019, doi: 10.1016/j.knosys.2019.05.028.

[31] M. B. Malik and D. H. Barus (Student Member, IEEE) was born in Medan, Indonesia. He received B.Eng (S.T) degree in electrical engineering and M.Eng (M.T) degree in power engineering from the Department of Electrical Engineering Institut Teknologi Bandung in 1998 and 2001, respectively. In 2002, he joined PT PLN (Persero), National Electricity Company in Indonesia. He is currently pursuing a Ph.D. degree in the Department of Electrical Engineering, Universitas Indonesia. He has authored or coauthored over 10 papers published in refereed journals and conferences. He has been a reviewer for IEEE Access. His research interests include power system operation and control, renewable energy, and AI applications in power system.

[32] D. H. Barus et al: Determining Optimal Operating Reserves toward Wind Power