

Research Article

Forecasting-Aided Monitoring for the Distribution System State Estimation

S. Carcangiu , A. Fanni , P. A. Pegoraro , G. Sias , and S. Sulis 

DIEE, University of Cagliari, Via Marengo 2, 09123 Cagliari, Italy

Correspondence should be addressed to G. Sias; giuliana.sias@diee.unica.it

Received 13 September 2019; Revised 13 December 2019; Accepted 11 January 2020; Published 28 February 2020

Guest Editor: Gonzalo Farias

Copyright © 2020 S. Carcangiu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, an innovative approach based on an artificial neural network (ANN) load forecasting model to improve the distribution system state estimation accuracy is proposed. High-quality pseudomeasurements are produced by a neural model fed with both exogenous and historical load information and applied in a realistic measurement scenario. Aggregated active and reactive powers of small or medium enterprises and residential loads are simultaneously predicted by a one-step ahead forecast. The correlation between the forecasted real and reactive power errors is duly kept into account in the definition of the estimator together with the uncertainty of the overall measurement chain. The beneficial effects of the ANN-based pseudomeasurements on the quality of the state estimation are demonstrated by simulations carried out on a small medium-voltage distribution grid.

1. Introduction

Power systems are rapidly evolving, and, correspondingly, control and management systems need information on the actual operating conditions of the grids, in order to manage them effectively. This information can be obtained in several ways and concerns different types of data, but, in any case, the proper operation of the systems is strictly connected to an accurate knowledge of the operating conditions of the grid [1]. Knowledge on possible operating conditions of the network can be obtained by means of load flow and power flow methods (see, for example, [2–4]), which are used, for instance, starting from nominal data in the planning stage or to find optimal configuration during contingencies. Other methods, suitable to know the state of the network, in terms of node voltages and/or branch currents, at run time starting from available measurements can be based on state estimation (SE) techniques. SE methods applied to the electrical power systems date back to the 1970s, when Fred Schweppe proposed the use of SE for achieving an accurate picture of the operating conditions of transmission networks [5]. The conventional state estimators assume the monitoring system to be overdetermined by having redundant measurements, thus ensuring the system observability, which is crucial for

the state estimator to work. Unlike transmission systems, the number of real-time measurements in distribution systems (DSs) is usually limited. Despite the recent vast deployment of smart meters (SMs), the monitoring systems of DSs are still underdetermined; moreover, possible metering or communication problems causing missing/delayed data contribute to real-time measurements scarcity. For DSs, SE is not feasible unless the so-called pseudomeasurements are introduced in order to ensure the observability of the system and to perform, thus, the so-called distribution system state estimation (DSSE). In reviews [6, 7], it is possible to find the context of research activity on DSSE techniques. In the following, the focus will be on the definition of pseudomeasurements for DSSE. Typically, pseudomeasurements are indirect or derived measurements used to describe load/generator power absorption/generation. It is worth recalling that loads in electric DS have highly different location, size, and typology, and distributed generation can be highly variable. Thus, an accurate idea of the behaviour of the overall load configuration is challenging and can be essential for an effective DSSE [8]. In order to allow DSSE to function properly, pseudomeasurements from load estimates and short-term load forecasts can be used [9, 10]. In recent times, research has been carried out on the application of machine

learning techniques both for load estimation and load forecasting problems.

Traditional load power estimation can be performed from typical load profiles of each user type. Assuming the demand at each hour is equal to the demand at the same hour of the previous equivalent day is a common choice. The load profiles can be considered as derived estimates of the customer load behaviour with high variance. Consequently, the quality of the state estimates obtained with this approach is generally poor. To improve the accuracy of load power consumption estimation in [11], a neural network approach is proposed to realign the average load profiles with the real power flow measurements available at the network substation. In particular, for each bus, two feed-forward artificial neural networks (ANNs) are trained, the first one associates real active power flow measurements with active power injections, whereas the second one relates real reactive power flow measurements with reactive power injections.

In the literature, algorithms based on Multiple Linear Regression (MLR) analysis [10] and machine learning techniques [12] have been proposed for load forecasting.

Load forecasting is a difficult task as the time evolution of the loads is complex and exhibits several levels of seasonality; the load at a given hour can be correlated not only with the load at the previous hour but also with the loads at the same hour on the previous day and with the same day in the previous week. Moreover, many important exogenous variables can be considered, especially weather-related variables. In this context, ANNs with their inherent capability to infer a function from data represent an efficient solution for load modelling more than linear models. Generating pseudomeasurements using ANNs for load forecasting has been demonstrated to be a viable solution especially for local-level load modelling [12]. In [13], a closed-loop estimator is proposed for a medium-voltage DS, where a machine learning function provides pseudomeasurements to DSSE. The output of the estimator is then fed back to the machine learning function, which allows the estimation when measurement data are missing. For each medium-voltage (MV) node, two feed-forward ANNs are trained to independently forecast nodal active and reactive power. Load time series and indices categorizing the load time series according to the load patterns are used as ANN inputs. The ANNs are retrained over time whenever a new load time series of an MV node is available. In [14], a SE tool with closed flow between the estimator and the machine learning function is proposed as in [13]. In [14], a feed-forward ANN is trained to forecast active power value one step ahead with three types of inputs: historical load information and weather-related and time-related variables. Once the predictive model has been developed, its performance is monitored continuously in order to detect the deterioration over the medium-to-long term (i.e., weeks to months) and retrain it accordingly.

Among all the proposed machine learning methods, multilayer perceptron (MLP) neural networks demonstrated to perform better than others for load forecasting [11, 15]. In particular, in [15], five machine learning methods, i.e., MLP,

Support Vector Machines, Radial Basis Functions, Decision trees, and Gaussian Process, have been compared to forecast the active power charges for the next 100 hours. The comparison has been made using three measures of accuracy (MAPE—Mean Absolute Percentage Error; MAE—Mean Absolute Error; and RMSE—Root Mean Squared Error) showing that the MLP is the most robust among the others.

In this context, this paper proposes an MLP load forecasting model for generating simultaneously high-quality active and reactive power pseudomeasurements for an effective branch-current-based DSSE (BC-DSSE). The estimator uses all the available information suitably considering uncertainty sources and correlations. In particular, appropriate weights are introduced to take into account the SM and forecast uncertainties. A realistic measurement scenario composed of few real measurements, SMs, load forecasts, and suitable pseudomeasurements is assumed. In particular, a feed-forward ANN is trained to predict one step ahead the power demands at each MV node. Exogenous variables have been used as predictive model inputs together with historical load information. A closed-loop information, flowing from the ANN outputs to the inputs, has been created to allow the BC-DSSE even though SM measurements are not available at the MV node for the last 24 hours.

The approach proposed in this paper is validated by means of simulation performed on a grid derived from a portion of a distribution network [16], which is a simplified version of an 18-bus UK radial feeder.

2. Distribution System State Estimation

DSSE is the key routine to obtain a picture of the network status at a given time instant. The system is locally considered under steady-state conditions, and the underlying measurement model can be described as follows:

$$\underline{z} = \underline{h}(\underline{x}) + \underline{\varepsilon}, \quad (1)$$

where \underline{z} is the vector of available measurements z_1, \dots, z_M ; \underline{x} is the vector including all the variables that uniquely define the state of the network; and $\underline{h}(\cdot)$ represents all the measurement functions (generally nonlinear) linking the reference measured values to the state of the system. Vector $\underline{\varepsilon}$ includes all the measurement errors and is considered a zero-mean random vector. The state \underline{x} can have different formulations (the so-called voltage state or current state, either in polar or rectangular coordinates), which are equivalent from a theoretical perspective but can lead to advantages in the implementation of DSSE solution routines. Hereafter, the following state \underline{x} is considered (see the BC-DSSE algorithm in [17]):

$$\underline{x} = \left[V_s \ I_1^r \ \cdots \ I_{N_{br}}^r \ I_1^x \ \cdots \ I_{N_{br}}^x \right], \quad (2)$$

where V_s is the voltage magnitude at a node of the network, which is chosen as the reference (e.g., the slack node), while I_k^r and I_k^x are the real and imaginary parts of k -th branch current, with the branch index ranging from 1 to the number of branches N_{br} . This branch-current formulation allows

linearizing some of the measurement functions and thus simplifying the estimation process.

From (1), it is clear that, to estimate the network state, it is necessary to manipulate the information coming from measurements in \underline{z} , also taking into account the characteristics of the measurement errors. As mentioned above, in DSs, it is hard to have a widespread installation of measurement devices and thus the availability of real-time measurements is typically limited. The measurement vector \underline{z} can thus be divided as follows:

$$\underline{z} = \begin{bmatrix} \underline{z}_{\text{real-time}} \\ \underline{z}_{\text{pseudo}} \end{bmatrix}, \quad (3)$$

where $\underline{z}_{\text{real-time}}$ includes the M_r real-time measurements, which can be voltage magnitude, current magnitude, and active or reactive power measurements as far as conventional measurements are concerned. Vector $\underline{z}_{\text{pseudo}}$ represents information that can be derived from other sources (pseudomeasurements), mainly historical information on active and reactive power consumption or generation. Every load or aggregated load of the network is thus analysed to define its average power absorption or injection. Pseudomeasurements are necessary to allow system observability in this context, but their accuracy is usually very low.

The most widespread technique to perform DSSE is represented by the Weighted Least Squares (WLS) approach, which aims at finding the state that minimizes the following objective function:

$$\hat{\underline{x}} = \underset{\underline{x}}{\text{argmin}} (\underline{z} - \underline{h}(\underline{x}))^T \cdot \underline{W} \cdot (\underline{z} - \underline{h}(\underline{x})), \quad (4)$$

which is the sum of the weighted squared residuals. Matrix \underline{W} represents the weighting matrix that allows penalizing or favouring each residual in a different way depending on the accuracy of the corresponding measurement or pseudomeasurement. \underline{W} is usually chosen as the inverse of the variance-covariance matrix of the measurements $\underline{\Sigma}_{\underline{z}}$ so that all measurements are weighted differently depending on their uncertainty (corresponding to a maximum likelihood estimation in case of normally distributed measurements), which is typically block-diagonal, since measurements performed by different devices can be considered uncorrelated.

The minimization in equation (4) is typically obtained by an iterative Newton solution of the following system of normal equations (considering the generic iteration i) [9]:

$$\underline{G}_i \Delta \hat{\underline{x}}_i = \underline{H}_i^T \underline{W} r_i = \underline{H}_i^T \underline{W} (\underline{z} - \underline{h}(\hat{\underline{x}}_i)), \quad (5)$$

where $\Delta \hat{\underline{x}}_i = \hat{\underline{x}}_{i+1} - \hat{\underline{x}}_i$ is the estimated state variation and $\underline{H}_i = d\underline{h}/d\underline{x}^T|_{\hat{\underline{x}}_i}$ is the Jacobian of the measurement functions in \underline{h} with respect to the state variables computed at the previously estimated state. The so-called gain matrix \underline{G}_i can be written as

$$\underline{G}_i = \underline{H}_i^T \underline{W} \underline{H}_i, \quad (6)$$

and it is constant when measurement functions are linear or can be linearized (see [17] for examples and details).

Pseudomeasurements are employed similarly to other measurements but are usually associated with very large

uncertainties (e.g., derived from historical variability of loads), which result in small contributions to the objective function in equation (4). Focusing on active power, a typical approach is to consider the following quantities for an unmonitored network node j :

$$z_{P_j} = \frac{1}{|\Gamma|} \sum_{n \in \Gamma} P_j(nT_s), \quad (7)$$

$$\sigma_{z_{P_j}}^2 = \frac{\sum_{n \in \Gamma} (P_j(nT_s) - z_{P_j})^2}{K},$$

where z_{P_j} and $\sigma_{z_{P_j}}$ indicate the active power pseudomeasurement and its standard deviation, respectively, $P_j(nT_s)$ is the active power injected (the injection convention is used for both absorbed and generated power) into node j , and T_s defines the time resolution of the available information (historical data). The set Γ includes the considered time instant indices (with n indicating the generic time instant) for all the available power samples, and $K = |\Gamma| - 1$ is its cardinality decreased by one to obtain the classical unbiased variance estimator. The higher the variability of the load/generator, the larger the corresponding variance, which is inversely proportional to the weight $w_{z_{P_j}}$ associated with the pseudomeasurement. A typical variation range of the power drawn by a load is over 50% of its nominal value, so it is easy to see that pseudomeasurements, while guaranteeing observability, are often of little help in improving the estimation accuracy.

In this context, SMs, and, in particular, those of 2nd generation, can play a significant role in enhancing pseudomeasurement definition. New SMs can provide voltage magnitude and active and reactive power measurements, with a much faster reporting rate than before. Italian authority for electric energy, for instance, gives and continuously updates directives on the functionalities for new-generation SMs, which include, among others, a 2 s measurement interval for “instantaneous” power measurements [18]. These measurements might be, in the future, directly acquired and integrated into estimation algorithms, but due to the huge number of installed SMs (above 30 million in Italy), it would be difficult to directly manage them in real-time and investment costs for communication and computation could easily become overwhelming.

For this reason, it is much more likely that SM measurements are used in an indirect and delayed way. The approach proposed in this paper is to collect active and reactive power measurements from SMs and exploit them to forecast the power consumption at a given time with anticipation compatible with the timing and data collection requirements. To this purpose, load forecasting Artificial Neural Network (ANN) models have been trained for different types of loads and/or aggregation of loads, as detailed in the following section. In fact, SM measurements from large customers or from a set of users connected to a given node (for instance, in a medium-voltage network) can be gathered from the field one day or few hours ahead the time instant of BC-DSSE execution and aggregated, so that they can serve as inputs in the neural load forecasting process.

As previously mentioned, once the forecast power consumption or generation of a given node is obtained, it can be included in \underline{z} as an enhanced pseudomeasurement and it is thus important to associate the correct covariance matrix to the new forecast quantities, thus allowing a correct weighting of the corresponding residual in the WLS procedure, according to equation (4).

Focusing on a generic bus j , the forecast procedure gives two predicted quantities P_j^F and Q_j^F , which are the active and reactive power injections, respectively. The following covariance matrix is thus needed:

$$\underline{\Sigma}_{\underline{z}^{PQ_j}} = \begin{bmatrix} \sigma_{z_{P_j}}^2 & \sigma_{z_{P_j}z_{Q_j}} \\ \sigma_{z_{P_j}z_{Q_j}} & \sigma_{z_{Q_j}}^2 \end{bmatrix}, \quad (8)$$

where $\sigma_{z_{P_j}}^2$ and $\sigma_{z_{Q_j}}^2$ are the variances of new pseudomeasurements z_{P_j} and z_{Q_j} , while $\sigma_{z_{P_j}z_{Q_j}} = \rho_{PQ_j} \sigma_{z_{P_j}} \sigma_{z_{Q_j}}$ is the covariance between the pseudomeasurement errors of the two forecast powers and ρ_{PQ_j} is the corresponding correlation coefficient. It is interesting to notice that, while in conventional DSSE approaches, the active and reactive powers pseudomeasurements are usually considered as uncorrelated, in the proposed approach, further information is taken into account and the correlation arising in the simultaneous estimation of P_j and Q_j can be easily included in the estimator by using the submatrix $\underline{W}_{\underline{z}^{PQ_j}} = \underline{\Sigma}_{\underline{z}^{PQ_j}}^{-1}$ in the overall weighting matrix.

Another important aspect, which is usually overlooked in the literature, is the modelling of the SM uncertainty. The proposed load forecasting is designed to obtain an estimation of the measured P_j and Q_j at a given time instant nT_s from previous available measurements, but this means that the computed power values can be considered only as approximations of the measured values (reference values are obviously unknown in practical conditions and real-time operation). The SM measurement chain is an additional source of uncertainty that affects the values considered in equation (8). As an example, the calibration process of SM devices cannot be perfect and compensate for all the systematic errors in all the operating conditions.

For this reason, in the following, the definition of the weights is discussed in detail in the presence of both forecast and SM errors. Focusing on z_{P_j} and z_{Q_j} , it is possible to distinguish the two zero-mean error contributions as follows:

$$z_{P_j} = P_j^F = P_j^{\text{ref}} + e_{P_j} = P_j^{\text{ref}} + e_{P_j}^F + e_{P_j}^{\text{SM}}, \quad (9)$$

$$z_{Q_j} = Q_j^F = Q_j^{\text{ref}} + e_{Q_j} = Q_j^{\text{ref}} + e_{Q_j}^F + e_{Q_j}^{\text{SM}}, \quad (10)$$

where P_j^{ref} and Q_j^{ref} are the ideal reference values of active and reactive power, $e_{P_j}^F$ and $e_{Q_j}^F$ are the corresponding forecast errors, and $e_{P_j}^{\text{SM}}$ and $e_{Q_j}^{\text{SM}}$ are the errors associated with the aggregated SM outputs. In the following, in the absence of further information, all the errors of the SMs associated with loads or generators grouped under the j th node are

considered independent. Similarly, the active and reactive power measurement errors are assumed uncorrelated.

Given the relative standard uncertainty α^{SM} of the generic SM (as derived from the SM datasheets and assumed as common to all SMs, without loss of generality), the relative standard uncertainty associated with $e_{P_j}^{\text{SM}}$ becomes

$$u_{e_{P_j}^{\text{SM}}} = \alpha^{\text{SM}} \frac{\sqrt{\sum_{i \in \Lambda_j} (P_i^{\text{SM}})^2}}{\sum_{i \in \Lambda_j} P_i^{\text{SM}}}, \quad (11)$$

where P_i^{SM} is the measured active power of load i , which belongs to the set Λ_j of the loads downstream the MV node j . Since the lack of knowledge in SM behaviour can be considered as independent from the prediction errors, the overall variance of the measurement z_{P_j} can be expressed as follows:

$$\begin{aligned} \sigma_{z_{P_j}}^2 = \sigma_{e_{P_j}}^2 = \sigma_{e_{P_j}^F}^2 + \sigma_{e_{P_j}^{\text{SM}}}^2 = \sigma_{e_{P_j}^F}^2 + u_{e_{P_j}^{\text{SM}}}^2 P_j^2 = \sigma_{e_{P_j}^F}^2 \\ + u_{e_{P_j}^{\text{SM}}}^2 \left(\sum_{i \in \Lambda_j} P_i^{\text{SM}} \right)^2. \end{aligned} \quad (12)$$

Similar expressions are valid for the reactive power while the correlation coefficient ρ_{PQ_j} , under the above assumptions, becomes as follows:

$$\rho_{PQ_j} = \frac{\sigma_{z_{P_j}z_{Q_j}}}{\sigma_{z_{P_j}} \sigma_{z_{Q_j}}} = \frac{\sigma_{e_{P_j}^F e_{Q_j}^F}}{\sigma_{z_{P_j}} \sigma_{z_{Q_j}}} = \frac{\rho_{PQ_j}^F}{\sqrt{1 + \left(\sigma_{e_{P_j}^{\text{SM}}}^2 / \sigma_{e_{P_j}^F}^2 \right)} \sqrt{1 + \left(\sigma_{e_{Q_j}^{\text{SM}}}^2 / \sigma_{e_{Q_j}^F}^2 \right)}}, \quad (13)$$

where $\rho_{PQ_j}^F$ is the correlation coefficient of the active and reactive power forecast errors.

The way these values are computed in practice is explained in the following section, where information available at each step is discussed. The above equations allow the computation of the elements of $\underline{\Sigma}_{\underline{z}^{PQ_j}}$ (see equation (8)) and thus, the definition of the weights for BC-DSSE that reflect appropriately the actual uncertainty in the proposed pseudomeasurement model.

3. Load Forecasting Neural Network Model

A Multilayer Perceptron (MLP) ANN with one hidden layer is used to forecast the load demand one step ahead (roughly speaking “ $t + 1$ prediction,” where the actual time step depends on the chosen ANN model in terms of input and output variables), where the time interval for prediction update is assumed equal to half an hour. In particular, residential and Small and Medium Enterprises (SMEs) loads have been considered in this paper. An MLP has been trained for each single or aggregated load. It has to be noted that, while for the residential loads, only the active power has been forecasted, as in the majority of the literature, for the SME loads, the corresponding reactive power is also considered in the proposed model. Figure 1 shows the structure of the MLP neural network for a SME.

The relationship between input and output patterns is described by the following algebraic equations system:

$$\begin{cases} \text{input layer} & \underline{W}_1 \cdot \underline{v} + \underline{b}_1 = \underline{y}, \\ \text{hidden layer} & h_l = f(\underline{y}), \\ \text{output layer} & \underline{W}_2 \cdot h_l + \underline{b}_2 = \underline{o}, \end{cases} \quad (14)$$

where \underline{v} is the input vector, which contains variables related to the time instant t :

- (i) Two weather variables at the time instant t : temperature, measured in °C, and humidity, in percentage;
- (ii) Three time-related variables: these consist of a label corresponding to the hour of the day [1, 2, . . . , 12], a label for the day of the week [1, 2, . . . , 7], and a label {1/0}, where 1 indicates a working day and 0 indicates a nonworking day;
- (iii) Four historical demand variables, which have a strong correlation with the recorded demand profile (active power for residential loads, whereas both active and reactive powers for SMEs): the demand at the previous hour, the demand at the same hour of the previous day, the demand at the same hour of the previous week, and the 24-hour average power evaluated considering all the recorded values in the previous day. As the considered time step is half an hour, the variables at the previous one hour, one day, and one week correspond to 2 steps, 48 steps, and 336 steps back.

The output \underline{o} of the network consists in the one-step ahead forecast load demand.

At time instant nT_s that corresponds to the current time tag of DSSE computation update, information on active and/or reactive powers of node j is needed. In the proposed solution, if the ANN forecast model is available for node j , powers P_j and/or Q_j are obtained, thanks to the performed prediction. In this case, the instant indicated as “ $t + 1$ ” in the forecast model description (see Figure 1) corresponds to the current time instant nT_s . For ease of presentation, a perfect match has been here adopted between T_s and the time interval of prediction update (half-hour step for both prediction and estimation updates), but other solutions at different rates are also possible following a similar scheme. As mentioned above, an ANN model is built for all the loads of interest and its outputs (load predicted powers) are fed into the DSSE algorithm at the following time step, with the procedure described in the previous section.

\underline{W}_1 is the weight matrix of the input layer, \underline{b}_1 is the bias vector of the input layer, \underline{y} is the input of the hidden layer, h_l is the output of the hidden layer, $f(\cdot)$ is the hidden neuron (logistic) activation function, \underline{W}_2 is the weight matrix of the output layer, and \underline{b}_2 is the bias vector of the output layer.

The Levenberg–Marquardt algorithm [19], which combines the gradient descent method and the Gauss–Newton method, has been used for the MLP training. The hyperbolic tangent sigmoid transfer function is used in the hidden layer. The inputs and outputs are normalized in the range $[-1, 1]$

before being used to train the ANN to balance the importance of input variables.

In the following, the case study and the database used to train and test the forecasting models are described.

4. Case Study

To evaluate the performance of the proposed approach, several tests have been carried out on a single-phase 18-bus network derived from a UK network (Figure 2) [20]. Connected to the 33 kV at bus 1, the network has a common rated bus voltage level at 11 kV. This network is used since it is adopted for other studies in the literature and gives a realistic load scenario for a MV network. On this topology, both industrial and residential loads have been considered as explained in detail in the next section. There is no loss of generality in considering loads information (SM measurements) coming from the database described in the next section, since individual loads are aggregated to replicate a load scenario that is compatible with nominal data of the considered network in [20]. Main assumptions and test results are also reported and discussed in the following.

4.1. Database for the Load Forecasting. To train and test the neural networks used to perform the load forecasting, data related to the active power consumption, available from the Commission for Energy Regulation (CER) [21], have been used. This database is anonym, and it consists of recorded half-hourly SM energy consumption from 6445 customers that participated in the “Electricity Smart Metering Customer Behavior Trials” [22]. The data are collected over a period of 18 months (from July 14, 2009, to December 31, 2010), at various distribution network locations in Ireland. The customer types are classified as residential (4225), SME (485), and others (1735). The present paper focuses on residential (the largest group in the available database) and SME customers who completed the trial. The SME customers are grouped into four subsectors: entertainment (including hotels, restaurants, sporting facilities, and public houses), industrial manufacturing, offices, and retail premises.

After removing the consumers having missing data, a database of 3423 residential and 287 SME consumers has been obtained. For SME loads, the reactive power profiles have been obtained starting from a real power factor (PF) profile recorded on a typical industrial site. In particular, the recorded PF profile has been propagated on the whole observation period. No reactive power was taken into account for residential loads.

Analyzing residential and SME data, a significant difference between the two consumer types can be observed. Figure 3 shows the power consumption (P , where the node index is dropped in the following when unnecessary, for the sake of simplicity) of four residential loads, randomly selected from the database. As can be noted, the energy consumption of each household is low, not regular, and very different in two consecutive working days as it depends on the lifestyle of its residents. On the contrary, the SME loads

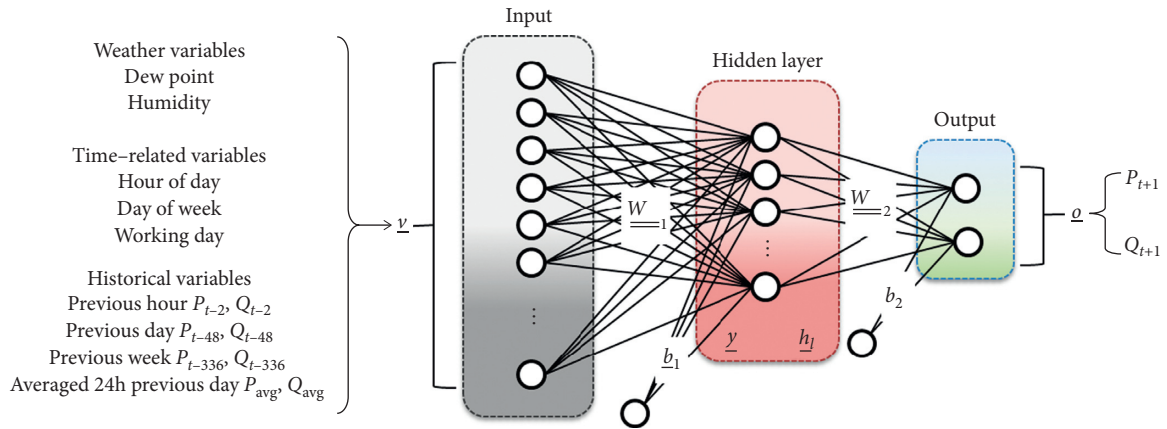


FIGURE 1: Structure of the active (P) and reactive (Q) power one-step ahead forecasting neural network model for SMEs (the structure of the neural network for residential loads considers only active power both in the input and the output layers).

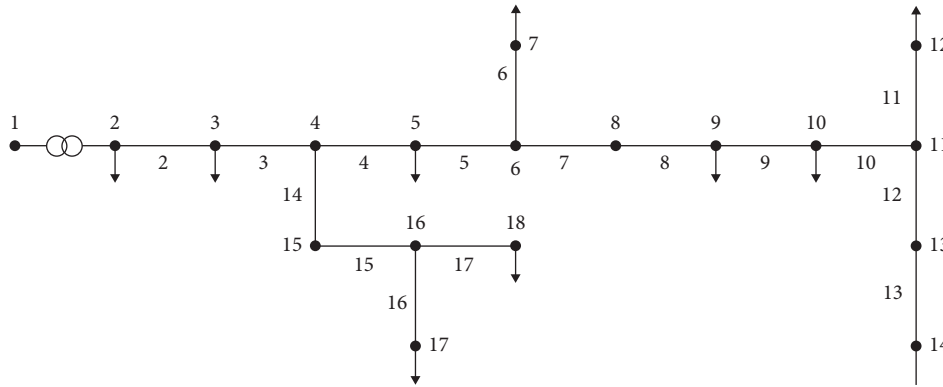


FIGURE 2: 18-node test network.

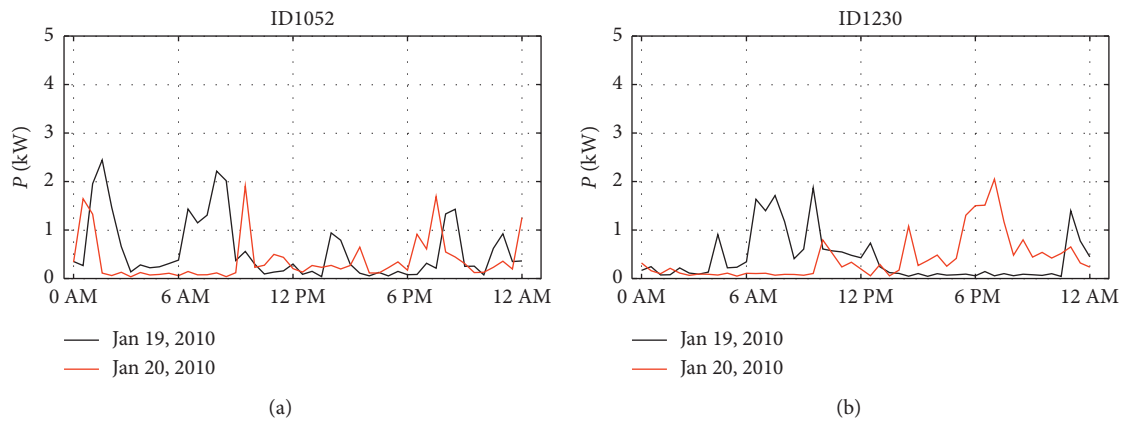


FIGURE 3: Continued.

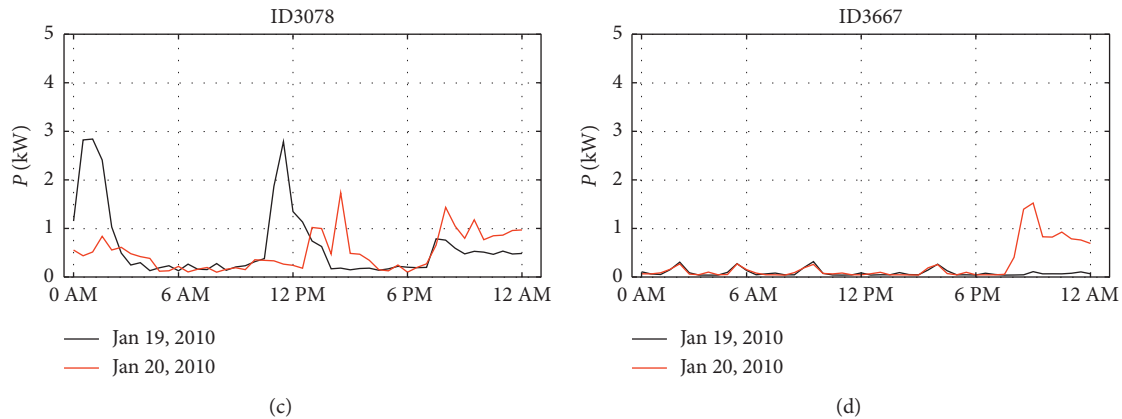


FIGURE 3: Power consumption of four residential loads in two consecutive working days.

(Figure 4) appear typically high and mostly regular due to the regular activity during the working hours and working days. Therefore, a different load forecasting performance for these two subsets is expected.

Since the objective of this paper is an effective state estimation of a DS, different equivalent MV loads have been created aggregating loads from the database with the aim to obtain power levels compatible with those of the considered grid (Figure 2). The load aggregation has been performed simply summing together the readings of several SMs. In particular, three different SME loads (sme_1 , sme_2 , and sme_3) are obtained by aggregating 21, 26, and 46 individual SMEs data, respectively. Moreover, two different residential loads (res_1 and res_2) are obtained by aggregating, respectively, the energy consumption of 523 and 537 residential loads randomly chosen. In Table 1, the ranges of the active power of the aggregate loads are reported.

Figure 5 shows the power consumption of one aggregated residential (res_1) and one aggregated SME (sme_3) load in the same two consecutive days considered in Figures 3 and 4. As expected, the aggregation makes the load profile more predictable. In fact, especially in the case of residential loads (Figure 5(a)), the aggregate load is more periodic and smoother than individual ones. This is because the aggregation operation permits to remove the high-frequency impulses corresponding to random events in the individual curve, alleviating and smoothing the randomness. Regarding the aggregated SME load (Figure 5(b)), its periodicity is more evident because the individual SME loads are already more periodic than individual residential loads.

The demand profiles depend not only on historical load evolution but also on exogenous variables, such as season and weather-related variables. Therefore, weather data, collected from the Irish Meteorological Service [23], have been added to the database. Among the weather variables, in this paper, the temperature (in $^{\circ}\text{C}$) and the humidity (in percentage) have been chosen as neural model inputs. This is because when the value of the temperature varies, the power system demand also varies. Furthermore, the humidity plays a relevant role in driving electricity demand during the warm months. In fact, the temperature above certain values is intensified by high humidity [24]. As the aggregation has

been performed by selecting individual loads located in different Irish areas, the simple averages of the temperature and the humidity percentage measured by several weather stations located in the central area of Ireland are used to represent the corresponding weather variables.

To highlight the dependence of the load energy consumption on the weather data, Figure 6 reports the daily energy consumption of res_1 and sme_3 and the average daily temperature trend over the same period. As can be noted, both residential and SME aggregated loads show a time pattern dependent on the temperature, with a stronger dependence for the residential load. Obviously, the dependence of the load energy consumption on the temperature data, even if prevailing, is not the only one. In fact, the drastic reduction in the consumption of SME loads and the increase in the residential one at the end of December 2009 are mainly related to the Christmas-New Year period rather than the temperature. In this paper, the dependence of the load energy consumption to the weather data is demonstrated by evaluating how the performance of the forecasting model is affected when the weather variables are excluded from the inputs of the model. The results are shown in the following section. Other weather variables, such as precipitation and wind speed, were analyzed, but it was found that, in this case, they do not have a significant impact on the energy demand.

As Figure 6 highlights, the considered energy consumption time series shows strong regularity, and a spectrum analysis revealed a prevalent daily periodicity, but at the same time, it is decidedly nonlinear. This complex data behavior can be captured by a MLP, which owns the ability to construct a larger set of nonlinear input/output mapping, by combining an appropriate number of nonlinear activation functions. Therefore, in this case study, a simple MLP, with a suitable structure, can be enough to build a performing forecasting model avoiding the overfitting of the training data.

4.2. Performance of the Load Forecasting Models. A neural load forecasting model has been designed for each aggregated load characterized by the range reported in Table 1. In order

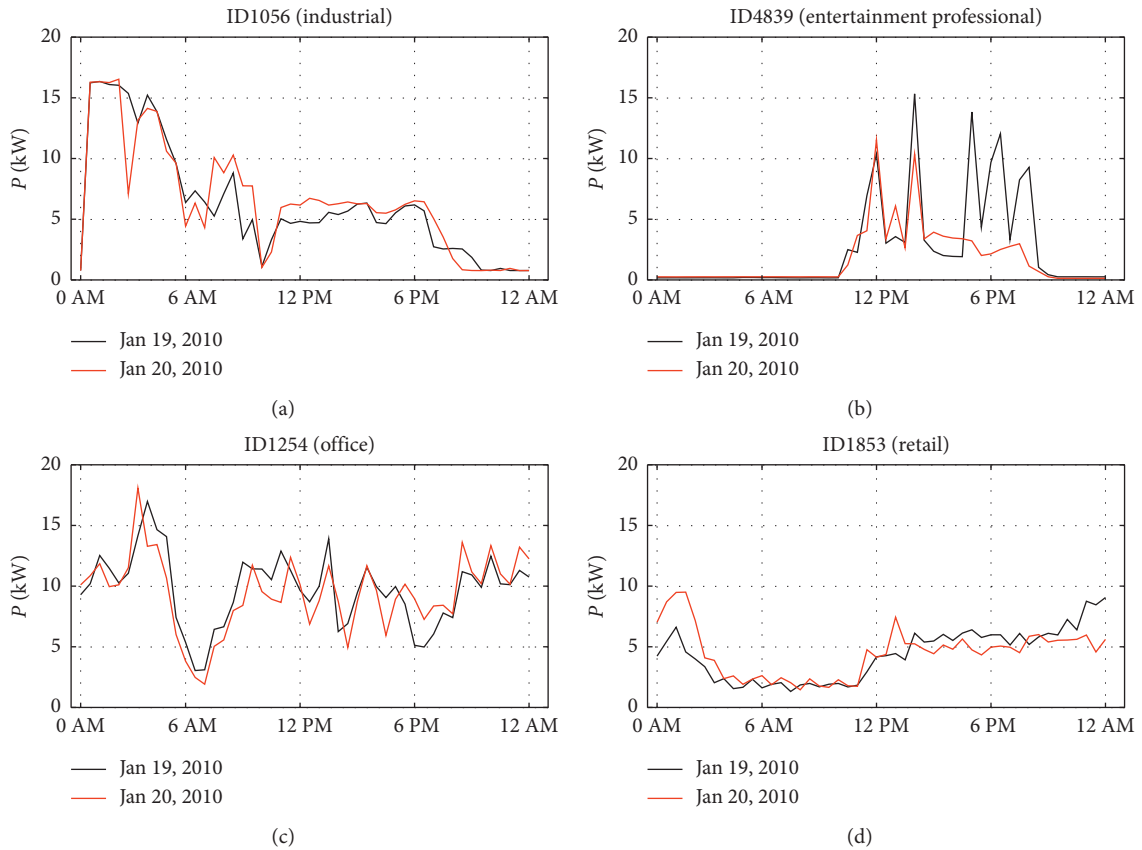


FIGURE 4: Power consumption of four SME loads in two consecutive working days.

TABLE 1: Range of active power P of the aggregate loads.

	sme_1	sme_2	sme_3	res_1	res_2
Min (kW)	67.70	101.64	192.32	69.68	73.05
Max (kW)	420.55	537.83	907.10	639.99	663.22

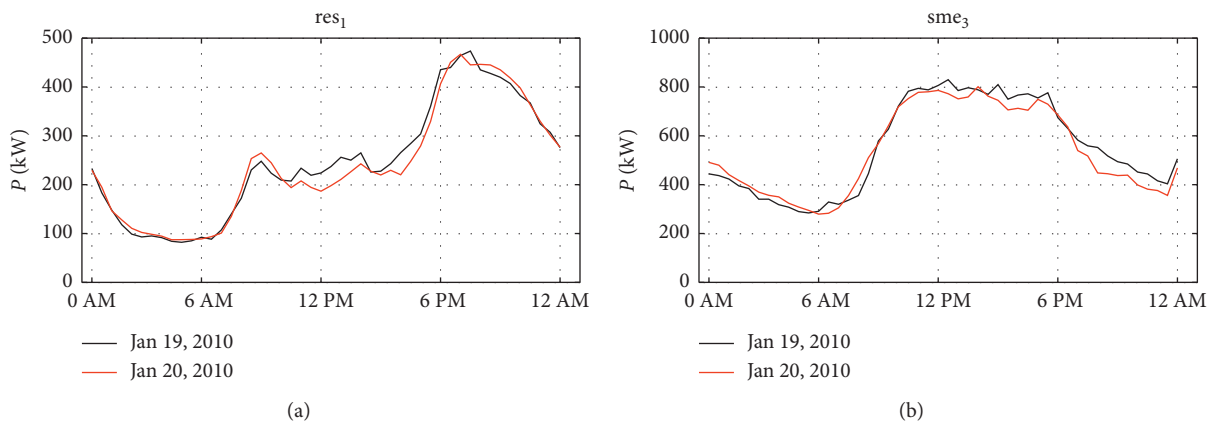


FIGURE 5: Aggregation of residential (a) and SME (b) loads.

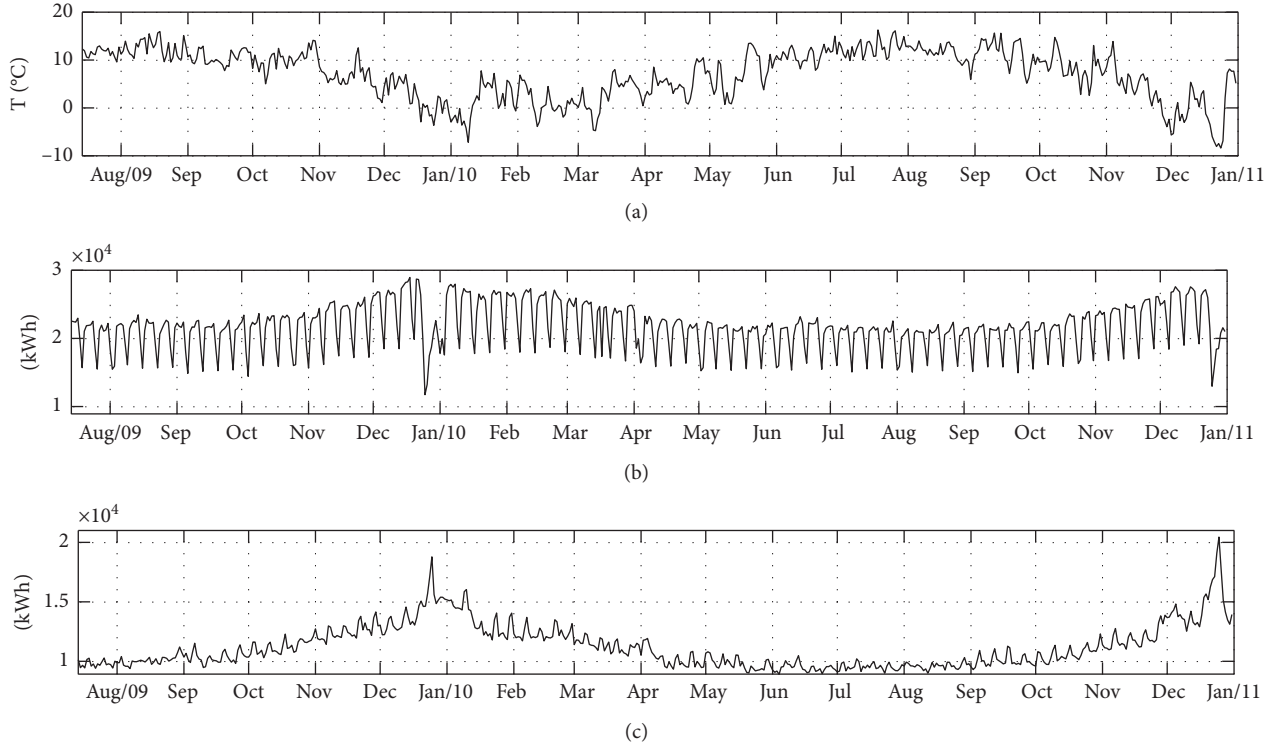


FIGURE 6: Residential and SME aggregate daily energy consumption compared with the average daily temperature trend. (a) Average daily temperature. (b) sme_3 daily energy consumption. (c) res_1 daily energy consumption.

to optimize the load forecasting network architecture, a trial-and-error approach has been performed to choose the appropriate number of hidden layer nodes, which consists in progressively growing the number of nodes, and selecting the network that minimizes the prediction error on the validation set. This optimization procedure resulted in 20 neurons for all the five networks (associated with loads in Table 1). Therefore, the best MLP architecture consists of an input layer with one neuron for each input variable (thus 9 or 13, for residential or SME, respectively), one hidden layer with 20 neurons, and an output layer with one neuron for each output variable (1 or 2 for residential or SME, respectively). Thus, the dimensions of the weight matrices and bias vectors in equation (14) result in 20×9 for \underline{W}_1 , 20×1 for \underline{b}_1 , 1×20 for \underline{W}_2 , and 1×1 for \underline{b}_2 in case of residential loads. In case of SME loads, \underline{W}_1 and \underline{b}_1 are 20×13 and 20×1 , respectively, whereas \underline{W}_2 and \underline{b}_2 are 2×20 and 2×1 , respectively.

The time series of each load profile is composed of 25728 half-hourly active and reactive (when considered) power values, from August 1, 2009, to December 31, 2010, while the July 2009 data were not used as only the recordings of fifteen days were available. The MLP training has been performed using the data of the first 12 months. The validation has been performed using the following 2 months (from August 1, 2010, to September 30, 2010). The last 3 months (from October 1, 2010, to December 31, 2010) have been used to test the trained neural model.

Since the forecasting accuracy depends both on the quality and quantity of the historical data used to train the predictor, a greater amount of data, for example, an extra year would certainly improve the prediction performance.

Note that a realistic assumption about the monitoring architecture could be that data from SMs are actually available after 24 hours. However, the proposed solution can also work with data collected one hour before (that could represent a future-proof scenario). Thus, during the training of the neural model, this information, both for active and reactive powers, has been included among inputs, because it is always available in the offline phase. In the online test phase, the corresponding inputs have been replaced by the values forecasted by the neural predictor at the previous steps. The outputs of the predictor are then fed back to the input layer creating a closed-loop information flow. This allows the state estimation even when measurement data are missing.

To evaluate the performance of the predictive models, the MAE, the MAPE, and the Root Mean Square Percentage Error (RMSPE), defined as in the following, have been used:

$$\begin{aligned} \text{MAE} &= \frac{1}{n} \sum_{i=0}^n |o_i - \hat{o}_i|, \\ \text{MAPE} &= \frac{100}{n} \sum_{i=0}^n \left| \frac{o_i - \hat{o}_i}{o_i} \right|, \\ \text{RMSPE} &= 100 \sqrt{\frac{1}{n} \sum_{i=0}^n \left| \frac{o_i - \hat{o}_i}{o_i} \right|^2}, \end{aligned} \quad (15)$$

where o_i is the actual load value, which can thus represent either the active power P or the reactive power Q of the considered load (respectively, P_j and Q_j when referring to

the network nodes as in the DSSE section above); \hat{o}_i is the corresponding predicted load value; and n is the number of training or testing samples. The smaller the values of MAE, MAPE, and RMSPE are, the better the forecasting performance is.

Figure 7 shows (in the top) the actual (black line) and the predicted (red line) active power load time series and (in the bottom) the corresponding differences between predicted and actual load powers for a month (October) of the test set related to sme_1 . Figure 8 reports the same time series for res_1 . As can be noted, the trends of the two loads are efficiently modeled by the neural predictors.

Figures 9 and 10 report the behavior of the actual and predicted real power load time series (top) and the corresponding prediction error (bottom), for the sme_1 and res_1 respectively, during the first test week. The MAPE for the forecasted active power in this time window results in 4.7% for the sme_1 and 4.8% for res_1 . Moreover, the validity of the zero-mean hypothesis has been verified for errors of both active and reactive power prediction.

In Table 2 and in Table 3, the training and test performances, obtained for the SME and residential aggregate loads, are reported, respectively.

It can be noted that, as expected, the performance deteriorates in the test phase. The SME loads show that the error percentages of the active and reactive power are very similar. The correlation coefficients between the real and reactive power errors are then evaluated to be used in the following, for the state estimation.

The results on the test set show that the proposed predictive model is able to forecast simultaneously both active and reactive powers (when required) with limited errors, starting from both exogenous and historical measurements. Moreover, it overcomes the problem of limited or time-delayed historical measurement availability throughout a closed loop information flow, which replaces the missing data with values forecasted by the predictor itself at the previous step.

The influence of the input variables on the load energy consumption can be evaluated through the performance of the forecasting model. In fact, the performance of the neural network model is expected to deteriorate when an effective variable is excluded from the inputs. In this paper, since aggregated loads show a time pattern dependent on the temperature, highlighted in Figure 6, the dependence of the load energy consumption on the weather data has been assessed. Removing the weather variables, the RMSPE on the test set increases by about 2%, for both active and reactive powers for SME costumers, and by about 1% for the residential ones (which are significant variations with respect to the results reported in Tables 2 and 3).

5. Performance of the Distribution Systems State Estimation

To assess the estimator performance, several simulations have been carried out starting from a measurement scenario that is realistic for a distribution grid. Two measurement points have been assumed on the network: on bus 1, with a

magnitude voltage measurement and an active and reactive power flow measurement; on bus 4, with a magnitude voltage measurement. SMs have been considered providing data fully available the day after the measurements. An accuracy equal to 1% for the magnitude of the voltage and equal to 3% for the power flows and SM measurements have been assumed.

A Monte Carlo approach has been applied in order to obtain statistically sound results, and the following assumptions are made:

- (i) Number of Monte Carlo trials, $N_{MC} = 1000$
- (ii) A maximum deviation of 50% with respect to the nominal values for the active and reactive powers drawn by the loads (uniform distribution)
- (iii) Measurement errors uniformly distributed

The SME loads sme_1 , sme_2 , and sme_3 have been connected to buses 17, 14, and 7, respectively (and thus are described by the couples $P_{17} - Q_{17}$, $P_{14} - Q_{14}$, and $P_7 - Q_7$), while residential loads res_1 and res_2 have been associated with buses 3 and 12 (and thus associated with P_3 and P_{12}). As for these loads, the last 1000 values of the test set have been considered, which correspond to a temporal interval of about 21 days. For each instant, a different operating condition of the network is thus considered by using such values for SME and residential loads and extracting the reference values applying the SM uncertainty. For all the other loads, active and reactive powers are extracted from nominal values according to the above assumption. Then, all reference values are computed from these load conditions by means of load flow calculation. Finally, measurements are also extracted from their random distribution and used as inputs to the BC-DSSE.

To assess the performance of the estimator, two different formulations and configurations of the BC-DSSE that correspond to different computation and management of the pseudomeasurements have been adopted. The first one, which uses the proposed estimator, exploits the predictions of the loads coming from the corresponding neural load forecasting models. This case is indicated as “*Prediction*” in Figures 11–13 reporting DSSE results in the following.

As mentioned in the previous section, the forecast active and reactive powers of SME loads (P_j^F and Q_j^F , with $j \in \{7, 14, 17\}$) have been used for each instant together with the forecast active power of residential loads (P_j^F with $j \in \{3, 12\}$).

To build the weighting matrix, besides the real-time measurement weights, submatrices concerning forecast loads must be included (see equations (9) and (10)) in the BC-DSSE. According to equation (12) (and its counterpart for reactive power), the variance of the pseudomeasurement is computed by using the RMSPE of the training set for the forecast errors and the datasheet information for the SMs. Since the SM measurements are not available in real time, the relative uncertainty of the aggregated power is evaluated for each time instant as the relative uncertainty of the day before at the same hour and it is

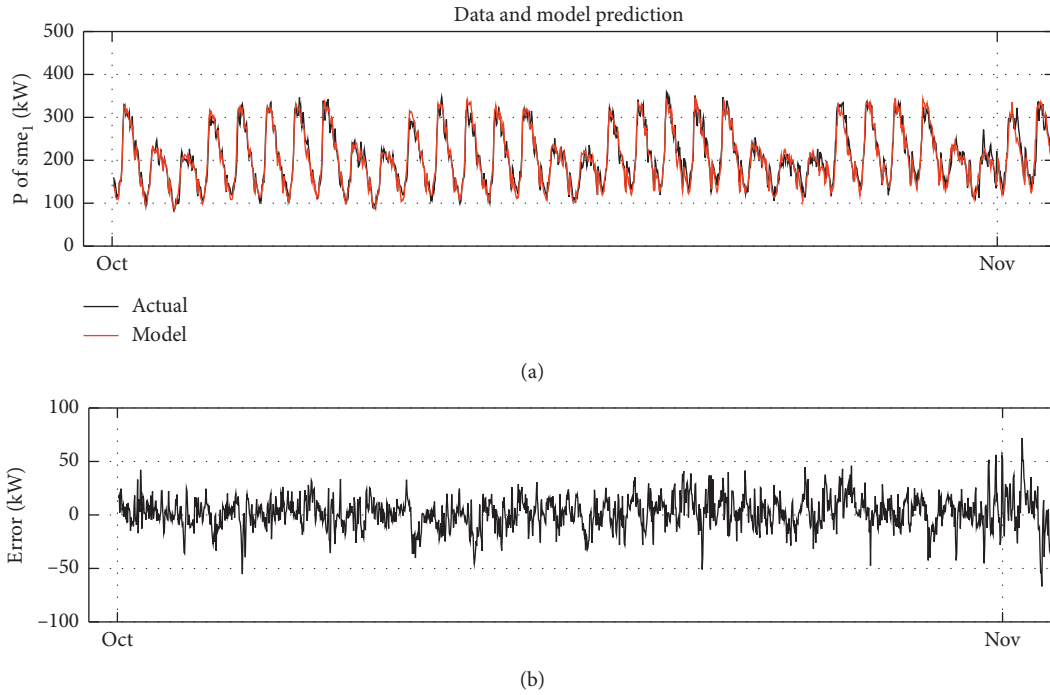


FIGURE 7: Test set behavior related to sme_1 : (a) the red curve represents the predicted real power values and black curve represents the actual one; (b) the differences between predicted and actual load powers.

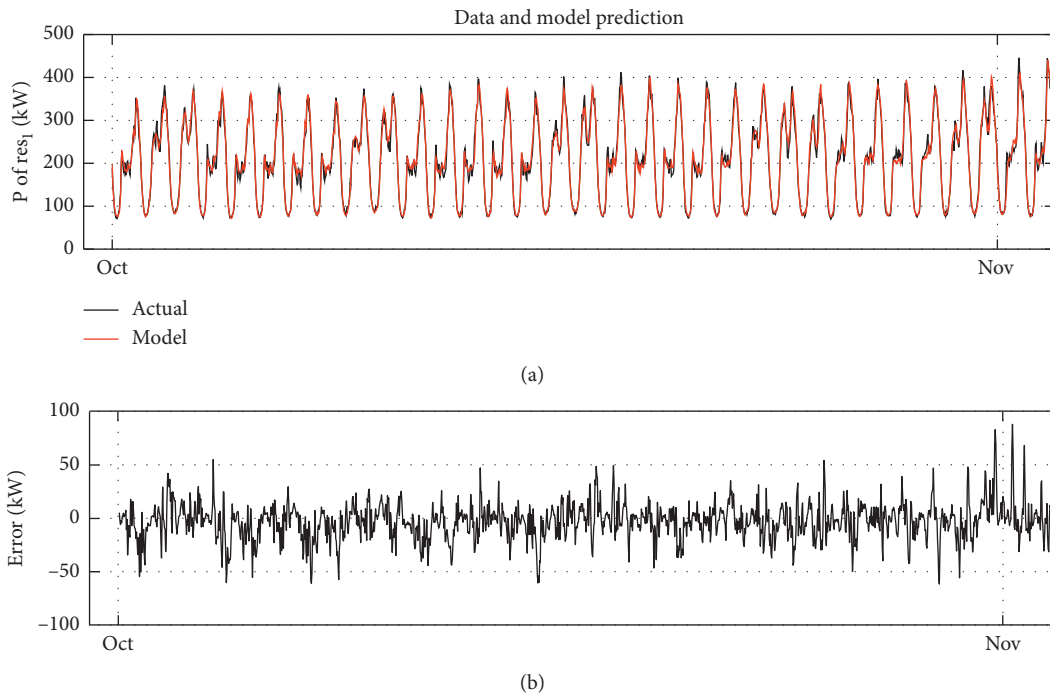


FIGURE 8: Test set behavior related to res_1 : (a) the red curve represents the predicted real power values and black curve represents the actual one; (b) the differences between predicted and actual load powers.

associated with the aggregated forecast powers at the current instant. The above procedure has been applied for all the estimations. It is interesting to notice that, with this

model, RMSPE can also be updated at fixed intervals by considering the measurements and forecast data obtained in the meanwhile.

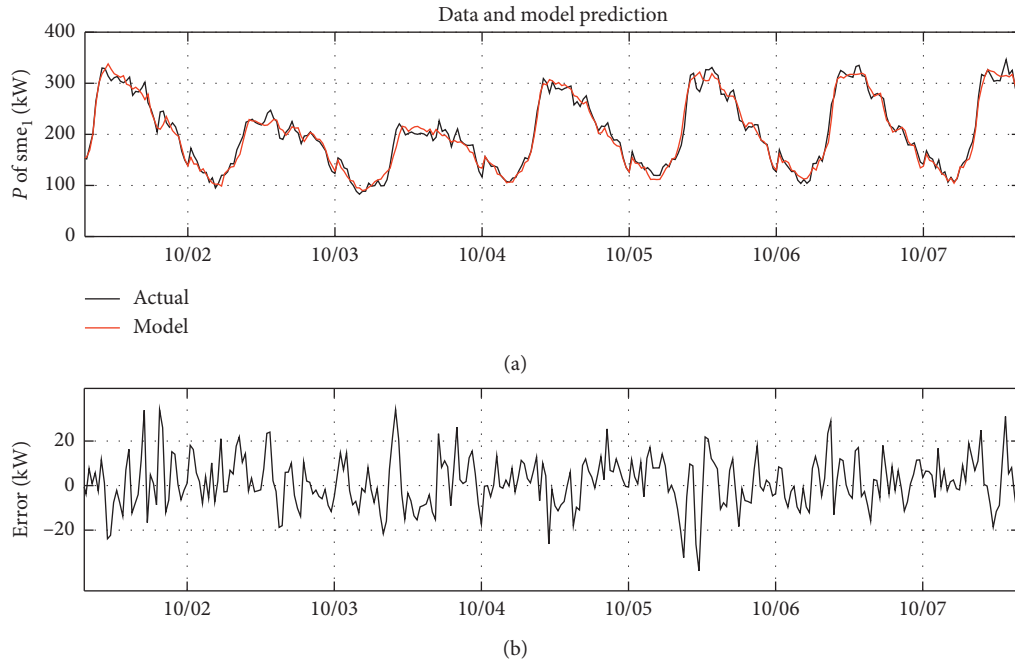


FIGURE 9: Detail of the first test week related to sme_1 ; (a) the red curve represents the forecasted real power values, and the black curve represents the actual one; (b) the corresponding error in kW.

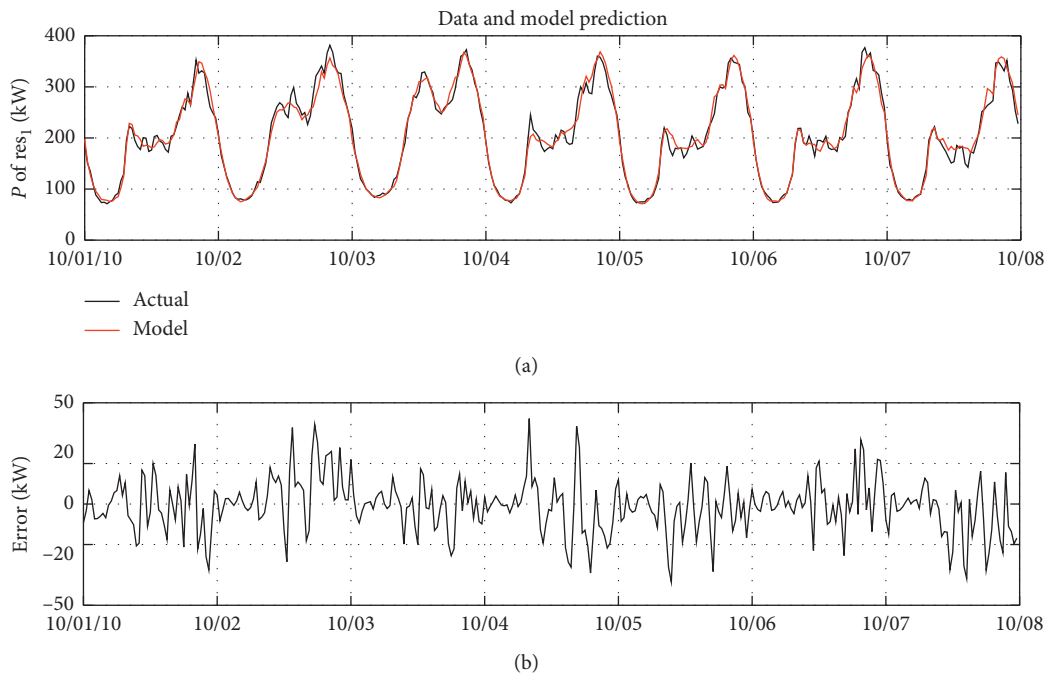


FIGURE 10: Detail of the first test week related to res_1 ; (a) the red curve represents the forecasted real power values, and the black curve represents the actual one; (b) the corresponding error in kW.

The second formulation corresponds to the classical BC-DSSE where no forecasting is considered, and pseudo-measurements of nodes 3, 7, 12, 14, and 17 are directly computed from the available measurement data. In particular, the measured power values collected the day before

at the same hour are used as pseudomeasurements. This estimator thus does not apply predictions (and is referred to as "No prediction" algorithm) and is considered as a benchmark for the proposed method in the same network scenarios.

TABLE 2: Performance of the neural load forecasting models for the SME aggregated loads (training and test).

Aggregate load	Active power (P)						Reactive power (Q)					
	MAE(kW)		MAPE(%)		RMSPE(%)		MAE(kvar)		MAPE(%)		RMSPE(%)	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
sme_1	11.2	15.2	5.5	7.4	7.5	12.0	2.1	2.8	5.5	7.4	7.5	11.9
sme_2	11.7	15.5	5.2	7.5	7.0	10.8	2.2	3.3	5.2	7.5	7.0	10.8
sme_3	17.4	26.9	4.1	6.2	5.4	9.1	3.3	5.1	4.1	6.2	5.4	9.0

TABLE 3: Performance of the neural load forecasting models for the residential aggregated loads (training and test).

Load	Active power (P)					
	MAE(kW)		MAPE(%)		RMSPE(%)	
	Train	Test	Train	Test	Train	Test
res_1	10.7	16.9	4.9	7.6	6.7	10.5
res_2	12.0	19.3	5.1	8.1	6.7	11.0

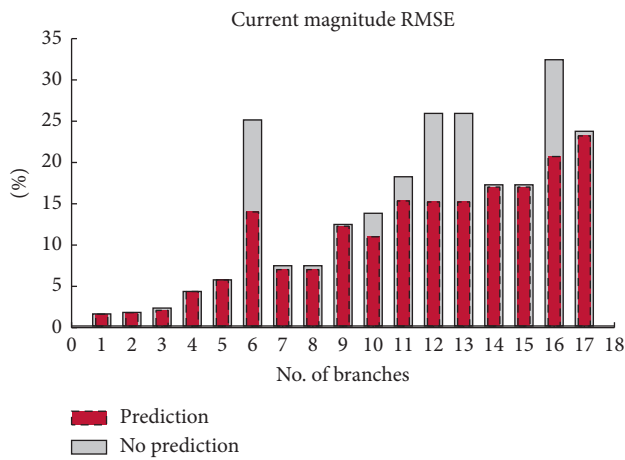


FIGURE 11: Current magnitude RMSEs obtained with and without the ANN-based prediction.

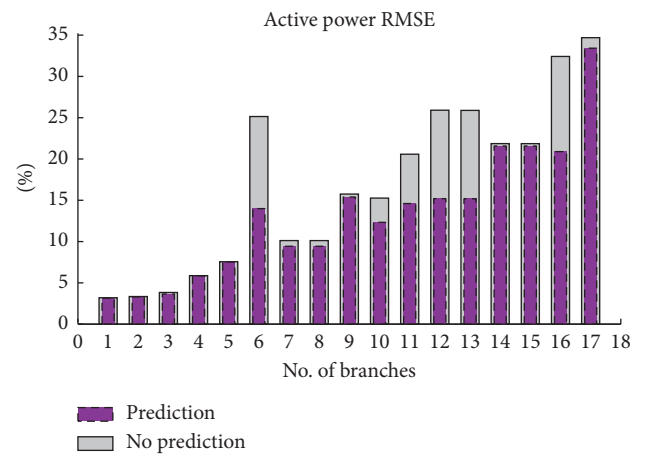


FIGURE 12: Active power RMSEs obtained with and without the ANN-based prediction.

Once the state variables (branch currents) are estimated along with derived quantities (e.g., voltages and power flows), a comparison of the results obtained with the two methods is performed in terms of percent root mean square errors (RMSEs) of the estimations (i.e., the square root of the mean of the squared differences between the estimated quantities and the corresponding reference values). RMSE results of the branch-current magnitude estimations are presented in Figure 11. The bar plot in red (dash line) shows the results obtained considering the prediction of the loads, while the bar plot in grey (the same holds for the Figures 12 and 13) presents the results obtained considering the above-described pseudomeasurements. A reduction close to 12% (meaning that the error is about halved) has been obtained as a best case on branch 6, where the largest forecast load sme_3 is connected, and an average reduction of more than 3% is also obtained. It is clear that the reductions in the estimation errors are more evident close to the position of the forecast loads. Branches 12 and 13 clearly show the same accuracy results, since node 13 is a zero-injection node. The same holds for the pairs 7-8 and 14-15.

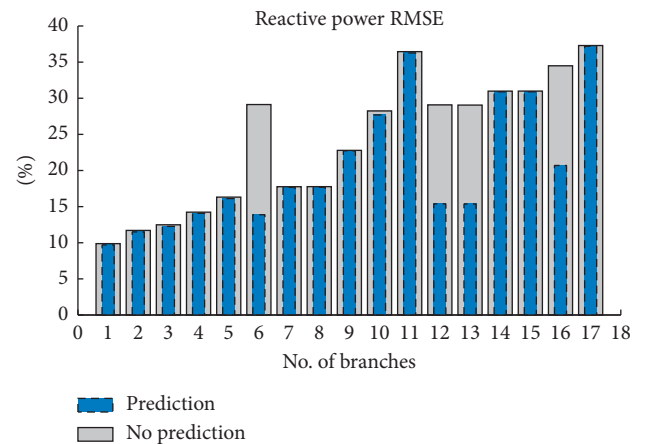


FIGURE 13: Reactive power RMSEs obtained with and without the ANN-based prediction.

Figure 12 shows the results obtained in terms of percent RMSE of the active power flow estimations for all the network branches. The bar plot in violet (dash line) shows

the results obtained considering the prediction. The considerations that can be drawn by these results are similar to those obtained for the branch currents: estimation improvements are more evident for the branches that are close to larger predicted loads. The error reduction is also more effective when lateral branches or leaves of the network are considered. In this case, the proposed algorithm brings a maximum reduction about 11.5% (error reduction of about 44%). Moreover, an average reduction of more than 3.3% is obtained.

As for reactive power estimations, it is possible to see in Figure 13 that the estimations are mainly affected by the prediction of the industrial loads locally, since the reactive power forecast is also available for them. A reduction of the percent RMSEs larger than 13% is obtained at branches 12, 13, and 16, while it is larger than 15% (the estimation error is more than halved in this case) for branch 6.

The test result highlights how the distribution state estimation performance significantly improves, introducing as pseudo-measurements the active and reactive power forecasted by the neural predictors instead of the power consumptions measured at the same hour of the day before. The improvements are more evident for branches close to larger loads.

6. Conclusions

Neural network load forecasting models demonstrated to produce reliable input information for a distribution state estimator, overcoming the problem of limited and time-delayed SM measurements or temporary failure in the communication system. In order to improve the accuracy of the state estimation, different requirements have been fulfilled: (i) the neural models are able to forecast simultaneously both active and reactive powers with limited errors, starting from both exogenous and historical measurements; (ii) the correlation between the forecasted real and reactive power errors has been determined, which results in significant information for the state estimation algorithm; (iii) a closed loop information flow allows the load forecasting, and hence, the state estimation, even when real measurement data are missing by replacing them with forecasted values; (iv) to build effectively the weighting matrix, needed to solve the state estimation algorithm, the variance of the pseudomeasurements can be updated at fixed intervals by considering the measurements and forecast data obtained in the meanwhile.

The test results show that introducing pseudomeasurements forecasted by the neural predictors significantly improves the DSSE and, more importantly, the improvements are more evident for the branches that are close to larger predicted loads almost halving the percent RMSEs of power and current estimations.

In summary, the proposed approach can be used for the state estimation of medium-voltage distribution networks that are either underdetermined, due to limited real-time measurements, or overdetermined but with delayed measurements from SMs.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research work has been supported by “Fondazione di Sardegna” within the research project “SUM²-GRIDS—Solutions by mUltidisciplinary approach for intelligent Monitoring and Management of power distribution GRIDS”—Triennial Agreement between Fondazione di Sardegna and Sardinian Universities, Regione Sardegna—L.R. 7/2007 year 2017—DGR 28/21, 17.05.2015. The article charge of this work was supported by the Open Access Publishing Fund of the University of Cagliari, with the funding of the Regione Autonoma della Sardegna L.R. n. 7/2007.

References

- [1] C. Muscas, M. Pau, P. A. Pegoraro, and S. Sulis, “Smart electric energy measurements in power distribution grids,” *IEEE Instrumentation & Measurement Magazine*, vol. 18, no. 1, pp. 17–21, 2015.
- [2] R. Madani, S. Sojoudi, and J. Lavaei, “Convex relaxation for optimal power flow problem: mesh networks,” *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 199–211, 2015.
- [3] S. Bolognani and S. Zampieri, “On the existence and linear approximation of the power flow solution in power distribution networks,” *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 163–172, 2016.
- [4] A. Adel Mohamed and B. Venkatesh, “Rectangular branch-based load flow,” *IET Generation, Transmission & Distribution*, vol. 13, no. 22, pp. 5151–5159, 2019.
- [5] F. Schweppe, “Power system static-state estimation, Part III: implementation,” *IEEE Transactions on Power Apparatus and Systems*, vol. 89, no. 1, pp. 130–135, 1970.
- [6] K. Dehghanpour, Z. Wang, J. Wang, Y. Yuan, and F. Bu, “A survey on state estimation techniques and challenges in smart distribution systems,” *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 2312–2322, 2019.
- [7] A. Primadianto and C.-N. Lu, “A review on distribution system state estimation,” *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 3875–3883, 2017.
- [8] D. Della Giustina, M. Pau, P. A. Pegoraro, F. Ponci, and S. Sulis, “Electrical distribution system state estimation: measurement issues and challenges,” *IEEE Instrumentation & Measurement Magazine*, vol. 17, no. 6, pp. 36–42, 2014.
- [9] A. Abur and A. Gomez-Exposito, *Power System State Estimation: Theory and Implementation*, Marcel Dekker, New York, NY, USA, 2004.
- [10] G. Gross and F. D. Galiana, “Short-term load forecasting,” *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1558–1573, Dec. 1987.
- [11] E. Manitsas, R. Singh, B. C. Pal, and G. Strbac, “Distribution system state estimation using an artificial neural network approach for pseudo measurement modelling,” *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1888–1896, 2013.

- [12] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 44–55, 2001.
- [13] J. Wu, Y. He, and N. Jenkins, "A robust state estimator for medium voltage distribution networks," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1008–1016, 2013.
- [14] B. P. Hayes, J. K. Gruber, and M. Prodanovic, "A closed-loop state estimation tool for MV network monitoring and operation," *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 2116–2125, 2015.
- [15] A. El khantach, M. Hamlich, and N. E. Belbounaguia, "Short-term load forecasting using machine learning and periodicity decomposition," *AIMS Energy*, vol. 7, no. 3, pp. 382–394, 2019.
- [16] A. R. Ahmadi and T. C. Green, "Optimal power flow for autonomous regional active network management system," in *Proceedings of the 2009 IEEE Power & Energy Society General Meeting*, pp. 1–7, Calgary, Canada, July 2009.
- [17] M. Pau, P. A. Pegoraro, and S. Sulis, "Efficient branch-current-based distribution system state estimation including synchronized measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 9, pp. 2419–2429, Sept. 2013.
- [18] ARERA, *Sistemi di smart metering di seconda generazione per la misura di energia elettrica in Bassa tensione*, ARERA, Milan, Italy, 2019.
- [19] J. Principe, N. R. Euliano, and W. Lefebvre, *Neural and Adaptive Systems*, John Wiley & Sons, Inc., Hoboken, NY, USA, 2000.
- [20] M. Pau, P. A. Pegoraro, and S. Sulis, "WLS distribution system state estimator based on voltages or branch-currents: accuracy and performance comparison," in *Proceedings of the 2013 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pp. 493–498, Minneapolis, MN, USA, May 2013.
- [21] Commission for Energy Regulation (CER), "CER smart metering project—electricity customer behaviour trial, 2009–201," in *Irish Social Science Data Archive: Belfield, Ireland*, Commission for Energy Regulation (CER), Dublin, Ireland, 1st edition, 2012, <http://www.ucd.ie/issda/CER-electricity>.
- [22] Electricity smart metering customer behaviour trials findings report," Tech. Rep. CER11080a, Commission for Energy Regulation, Dublin, Ireland, 2011.
- [23] Met Eireann (Irish Meteorological Service), <http://www.met.ie>.
- [24] J. Xie, Y. Chen, T. Hong, and T. D. Laing, "Relative humidity for load forecasting models," *IEEE Transactions on Smart Grid*, vol. 9, no. 1, pp. 191–198, 2018.