SHORT-TERM OPERATIONAL PLANNING AND STATE ESTIMATION IN POWER DISTRIBUTION NETWORKS

Barry HAYES
IMDEA Energy – Spain
barry.hayes@imdea.org

Milan PRODANOVIC
IMDEA Energy – Spain
milan.prodanovic@imdea.org

ABSTRACT

This paper discusses the application of a closed-loop state estimator for improving situational awareness in distribution systems. A predictive database is created and applied to forecast future network states, in order for short-term (e.g. day-ahead) planning to be carried out. This can be used to provide early warning of potential network issues and more optimal management of distributed energy resources. The predictive database is based on adaptive load forecasting models, which are continually updated based on feedback from the state estimator. The methodology is demonstrated using data recordings from an actual MV distribution network.

INTRODUCTION

Recent years have seen a large increase in the penetration of Distributed Energy Resources (DER) in MV and LV networks, including Distributed Generation (DG), demand responsive loads, and storage devices. This has resulted in a much higher level of variability and complexity in distribution network operation, and the need for better situational awareness and a more pro-active system support. Hence, there has been considerable interest in adopting a number of operational techniques, previously only used at the transmission level to distribution systems, such as state estimation and short-term operational planning [1]-[4]. This paper presents a methodology for short-term operational planning (e.g. hours/days-ahead analysis) of distribution systems with DER. The approach uses real-time and historical measurement data from LV smart meters and/or MV SCADA, and information on local weather conditions to develop a predictive database for forecasting demand and DG outputs at each network node. This allows load/DG to be forecasted, and the future states of the distribution network to be estimated using Distribution System State Estimation (DSSE).

The predictive database is continuously updated and improved based on feedback from the DSSE and the real-time measurements from the network. This creates a closed-loop information flow, which allows the network operator to estimate the distribution system state accurately, even when measurement data are missing (e.g. due to metering or communication errors), or in hours/days ahead analysis of the system for short-term operational planning. This approach can help to provide early warning of potential network issues and allow more optimal scheduling and management of DER. The analysis is carried out using a case study of the MV distribution network in [5], for which detailed

recordings of demand and generation are available at the end-user (home/factory level) and at the MV substations (10:0.4 kV transformers).

METHODOLOGY

Overview of Methodology

The input data to the system are static network parameters (bus and branch information), along with measurement data (e.g. real-time recordings of V, P, or Q at any bus/branch in the network), and/or pseudo-measurements (estimated or forecasted load demands and/or DER output). The approach is also designed so that the user can operate in "real-time" mode, applying real-time measurements from the network in the analysis, or in "forecasting mode" to carry out hours/days-ahead analysis of the system. In forecasting mode, all of the inputs to the system are pseudo-measurements, based on short-term forecasts of demand and DG outputs. These data are fed to a robust DSSE, which identifies bad data, such as erroneous or missing values in the input measurements. This allows network analysis, such as power flows and contingencies to be carried out accurately. Finally, in the post-processing phase, trending or out-of-control parameters are identified, and warnings, alarms and recommendations are provided to the network operator. One of the novel aspects of this approach is the use of feedback from the post-processing phase, which makes the DSSE a closed-loop system. This allows the predictive database and forecasting tools to be continually updated as more real-time data becomes available. This approach is in contrast to most DSSE methods in the literature which are open-loop. Fig. 1 shows a flow chart of the overall methodology.

Robust Distribution System State Estimation

State estimators (SEs) have been a standard feature of transmission network operation for many years, where they are applied to improve the observability of the network and reduce the impacts of noise and errors in system measurement data. Since the characteristics of distribution networks are fundamentally different from transmission networks (e.g. high R/X ratios, radial configurations, reduced quantity and quality of system measurements), many of the well-established techniques used in transmission SE cannot be applied directly. Recently, a number of SEs specifically intended for distribution systems have been proposed [1-2, 7-8].
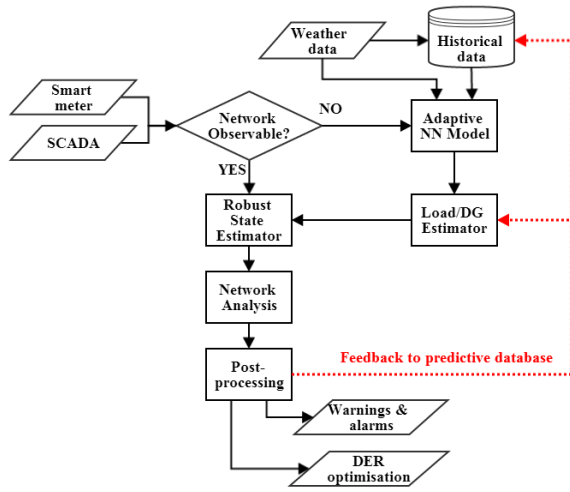
Figure 1 - Flowchart of methodology

Input Data and Observability

The observability of any power network depends on the number, type and locations of the available measurements. This can be determined by examining the null space of the network measurement Jacobian matrix according to the method in [6]. If the network, or any parts of it, cannot be observed, estimates of the demand and DG output are used to provide pseudo-measurements of power injections at each relevant network node. Therefore the DSSE may use various combinations of input data, comprising (in order of decreasing accuracy) of: (i) real-time measurements, (ii) load pseudo-measurements, and (iii) forecasts of future load/DER outputs. Each of these input data types potentially contains significant noise and gross errors. It was found that a least-squares estimator based on robust statistics [9] was required to produce acceptable performance in this application. Estimators based on robust statistics are particularly suited to dealing with gross errors and outlier values which can cause computational problems for conventional SEs.

DSSE Procedure

Most DSSE algorithms operate by minimizing the conventional Weighted Least Squares (WLS) objective function [7]:

$$min_x \qquad J(x) \ = \ \sum_{i=1}^{N} w_{ii} \, (z - h(x))^2 \qquad (1)$$

$$min_x \qquad (z - h(x))^T \overline{W}(z - h(x)) \qquad (2)$$

$$Subject \ to \qquad (z = h(x) - \varepsilon) = h(\hat{x}) - r \qquad (3)$$

where w_{ii} is the weight for measurement i, z is the input data vector, $h(x)$ are the measurement functions, x is the state vector, $\hat{x}$ is the estimated state vector, ε is the measurement error, $\overline{W}$ is the measurement weight matrix, and r is the vector of measurement residuals. The weights in the diagonal of $\overline{W}$ are set according to the

variance σ_i^2, of each meter (in the case of real measurements), or the variance of each load forecast (in the case of pseudo-measurements). This is so that the SE solution gives more weight to measurements which are known to have greater accuracy:

$$\overline{W} = \begin{bmatrix} 1/\sigma_1^2 & 0 & \cdots & 0 \\ 0 & 1/\sigma_2^2 & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\sigma_i^2 \end{bmatrix} \qquad (4)$$

Equation (2) is solved iteratively as follows:

$$\Delta z_n = \ z - h(x_n) \qquad (5)$$

$$\Delta x_n = (H^T \overline{W} H)^{-1} H^T \overline{W} \Delta z_n \qquad (6)$$

$$x_{n+1} = x_n + \Delta x_n \qquad (7)$$

where n is the number of iterations. However, it was found that in the presence of significant input data errors, the conventional WLS approach can have computational issues which result in the SE becoming insoluble. These problems are caused by poor conditioning of the network Jacobian matrix H, which results in difficulties inverting H to form the gain matrix $G = (H^T \overline{W} H)$. These issues were overcome by applying the equivalent weight function proposed in [2]. The diagonal measurement weights matrix $\overline{W}$ is modified as follows:

$$\overline{W} = diag(\overline{w}_1, \overline{w}_2, \dots, \overline{w}_N) \qquad (8)$$

where the equivalent weights $\overline{w}_i$ are re-calculated at each SE iteration:

$$\overline{w}_i = \begin{cases} w_i, & D_i' \leq k_0 \\ w_i \frac{k_0}{D_i'} \left(\frac{k_0 - D_i'}{k_1 - k_0} \right)^2, & k_0 < D_i' \leq k_1 \\ 0, & D_i' > k_1 \end{cases} \qquad (9)$$

where
$$D_i' = \frac{|r_i|}{(1 - J_{ii} + \beta)\theta} \qquad (10)$$

$$\theta = \alpha \, med_j |r_j - med_i \, r_i| \qquad (11)$$

$$k_0 = min \, (K_{t1}, K_{t2}), \quad k_1 = max(K_{t1}, K_{t2}) \qquad (12)$$

$$K_{t1} = \alpha \, med \, (D_i'), \quad K_{t2} = K_{t1} + (\max(D_i' - K_{t1}))/3 \qquad (13)$$

Also $i, j = 1,2,…,N$, $\alpha = 1.438$, β is a small constant added to avoid division by zero, and med is the median. This iterative re-calculation of the DSSE weights is based on the influence function from robust statistics theory [9]. It reduces the influence of measurements with extreme values which can cause the estimator to break down, by decreasing the weights $\overline{w}_i$ if D_i' approaches the upper threshold k_1. This inclusion of the equivalent weight

function was crucial in making the DSSE robust to gross errors, particularly when a large number of load estimates and forecasts are applied.

Adaptive Load Estimation and Forecasting

In order to provide suitable pseudo-measurements for the DSSE, the load demand (and DER output) at each measurement node in the network must be provided. High-quality pseudo-measurements enable the DSSE to function effectively, even in cases when there are measurement errors or missing data due communication system failures, and also for analysis in forecasting mode (e.g. day-ahead planning). First, the parameters which have a significant influence on demand were identified. The analysis identified 8 parameters: two weather-related parameters (temperature, dew point), three time-related (day, hour of day, and whether or not the day is a normal working day) and three historical load parameters (previous hour demand, previous week equivalent hour demand and previous 24 hour average).

Comparison of Approaches for Load Forecasting

The above data can be used to forecast demands using various techniques. The most common approaches are: (i) Multiple Linear Regression (MLR) analysis; (ii) using typical load curves for each user type (e.g. residential, industrial); and (iii) Neural Networks (NN) [8, 10]. Each of these techniques was applied to estimate the load at each of the MV substation nodes in the 10kV case study network [5]. In total, there were 43 loads, comprised of 30 purely residential loads, and 13 other loads which included factories, as well as some district heating and water pump loads (these are classified together as "industrial" for simplicity).

Table 1 shows the Mean Absolute Percentage Error (MAPE) obtained using each load forecasting technique, for the aggregate system load, and the average MAPE values obtained for all of the individual residential and industrial loads.

Case	Estimation Method	Aggregate system MAPE (%)	Residential Loads MAPE (%)	Industrial Loads MAPE (%)
(i)	MLR	9.40	15.38	68.84
(ii)	Load curve	8.82	14.71	42.07
(iii)	NN	5.34	11.58	35.17

Table 1 – Comparison of day-ahead load forecasting results.

From Table 1 it can be seen that all three techniques are more accurate at predicting the overall system aggregate demands than individual demands. This is as expected, since the individual loads are smaller (individual bus peak demands range from 20-400 kW, while the overall system aggregate peak is 3.2 MW). Therefore individual demands are affected much more by random load switching, and have higher variability. The above

analysis also demonstrates that non-linear methods of load forecasting had better performance (NN, case (iii)) than linear methods (cases (i)-(ii)). In some previous studies on load estimation, it has been noted that non-linear methods, such as NN, have relatively similar performance to more conventional linear methods [10]. However, most previous work in this field applies to much larger, more aggregated load, e.g. prediction of regional or national demands. In this paper, the focus is on local-level load estimation (e.g. for few tens to hundreds of residential customers). Here the relationships between demand and historical/weather data are much more non-linear, and conventional forecasting techniques do not perform well. For these reasons, a NN-based approach was selected.

Adaptive Load Estimation using Neural Networks

Autoregressive models are widely used in short-term load forecasting. These models combine the use of a static load estimation model trained with historical data, and recent load measurements from the network. Non-linear Auto-Regressive Exogenous (NARX) models are particularly useful for modelling load times series with non-linearities [11]. The NARX model is expressed as:

$$y_{t+1} = F(y_t, y_{t-1}, \dots, y_{t-do},\ u_t, u_{t-1}, \dots, u_{t-di}) \quad (14)$$

where the next value of the output signal (e.g. the MW load), y_{t+1}, is regressed using previous load measurement values y_t, y_{t-1} and the input signals u_t, u_{t-1} such as time and weather load variables. The variables *di* and *do* are the number of time delays in the input and output layers, respectively, which can be adjusted to allow for different forecasting horizons, e.g. hour-ahead, day-ahead etc. The function *F* represents the two-layer autoregressive NN illustrated in Fig. 2. The neural network weights are indicated by *w* and the autoregressive model coefficients are denoted *b*. The performance of the NARX model is shown in Table 2.
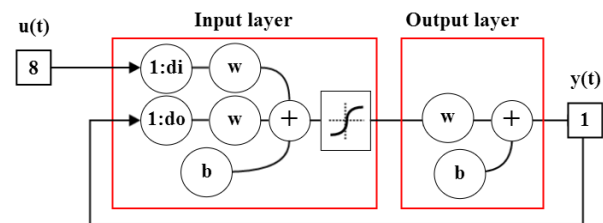

Figure 2 – Structure of NARX model.

Case	Estimation Method	Aggregate system MAPE (%)	Residential Loads MAPE (%)	Industrial Loads MAPE (%)
(iv)	NARX	3.30	8.70	27.79

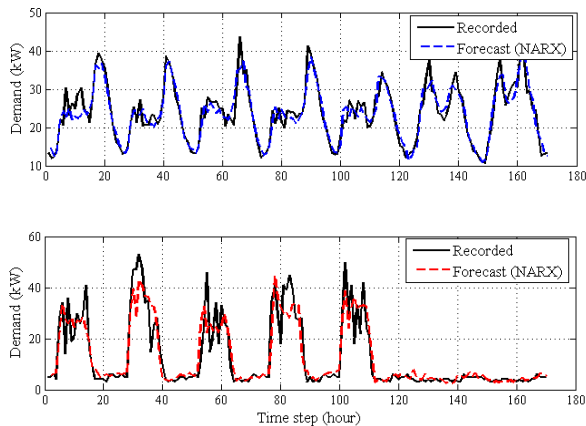**Table 2** – Day-ahead load forecasting results using NARX model.

Figure 3 – Samples of day-ahead demand forecasting results at residential and industrial MV/LV transformer substations.

Samples of the recorded and day-ahead forecasted MW demands for selected individual residential and industrial load buses are provided in Fig. 3. A significant advantage of the NARX model is that the model adapts to short-term changes in load behaviour, as the autoregressive coefficients are automatically updated based on recent measurements. For longer-term changes, e.g. seasonal changes, or changes in user consumption patterns due to a tariff change or DER installation at a particular bus, the model NNs should be re-trained. Underperforming load estimation/forecasting models can be identified and re-trained based on feedback from the DSSE (see also Results and Discussion). The average NN training time for each MV node was 1.3 seconds (using MatLab on a two-core, 2.6 Ghz Intel i3 microprocessor).

RESULTS AND DISCUSSION

The closed-loop DSSE described in the previous sections is applied to the case study distribution network for which detailed recordings are available at MV substations and at the end-user level in the LV network. The network is a 48-bus, 10kV system with a weakly-meshed structure, Fig. 4. The network has a peak demand of 3.2 MW, which is made up primarily of suburban/rural residential customers (77% of total demand). There is also significant PV generation installed at LV, and district heating load. At most nodes in the network, measurements of MW demand only were available. Hence, in order to calculate reactive powers, aggregate power factors of 0.97 for residential loads and 0.95 for industrial/other loads were assumed.

Performance of Closed-loop DSSE
Fig. 5 shows the performance of the DSSE, expressed as the MAPE of the voltage, MW, and MVAr flows throughout the distribution network. Several scenarios are tested. The initial conditions at time $t = 0$ are: input measurements of real and reactive power injections with standard deviation $\sigma = 5\%$ at each MV node, and a
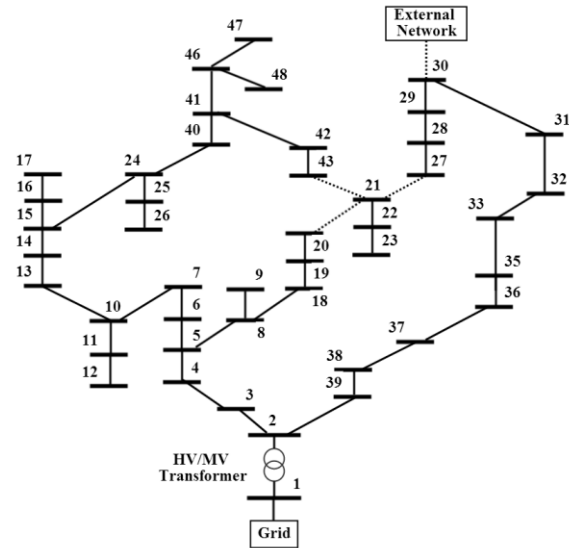

Figure 4 – MV distribution network case study [5].

measurement redundancy factor of 1.03 (slightly over-determined). The following scenarios A-E were simulated and some comments on the results from the DSSE are given in each case:

- **Scenario A**: Gross measurement error at single bus ($t = 100$). This results in a small error in voltage and real power flow.
- **Scenario B**: Multiple missing measurements from buses 46-48 ($t = 200$). The system is now under-determined (redundancy factor 0.97), and DSSE replaces the missing inputs with pseudo-measurements. There is only a minor impact on accuracy of estimation of voltages and flows.
- **Scenario C**: Complete failure of communication system ($t = 300$). All measurement inputs are replaced with load estimates from case (i) MLR. The DSSE continues to function, although some large errors occur (MAPE $\approx 60\%$).
- **Scenario D**: Complete failure of communication system ($t = 500$). All measurement inputs are replaced with load estimates from case (iv) NARX. Voltage and power flow estimation errors are significantly reduced compared to Scenario C.
- **Scenario E**: MV node injections replaced with more accurate measurements with $\sigma = 1\%$ from e.g. improved LV smart metering ($t = 700$).

Short-term Operational Planning
The load forecasting and robust DSSE can also be applied to short-term planning in the distribution network. For example, hours/days ahead contingency analysis can be carried out to estimate the impacts of unscheduled faults. This can provide the network operator with early warning of potential voltage or thermal line rating excursions, and recommendations for corrective actions (e.g. network re-configuration). This could also be used for optimal management of DER.
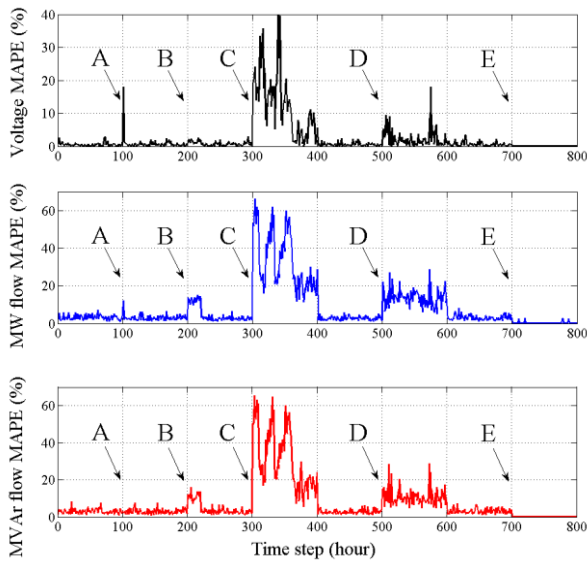
Figure 5 – Mean absolute percentage errors for system voltages, real and reactive power flows obtained with various DSSE inputs.

Applying a closed-loop DSSE approach allows for the design of operational planning services which can run automatically, with minimal intervention from the user. The NN load models have the advantage that they can be trained at any MV node without a priori knowledge of the load structure. These models can then adapt according to observed changes/trends in load behaviour. This makes the approach in Fig. 1 straightforward to apply to a wide range of distribution networks and load types.

Identification of Bad Data and Feedback to Load Estimator

The measurement residual vector *r* describes the "fit" of the measured input data to the DSSE equations (1)-(7). Statistical tests, such as the Chi-squared test, are used to determine if a given measurement is bad, i.e. a poor statistical match to the overall measurement data set and system equations. Individual bad data points are automatically removed by the DSSE.

In addition, a record is kept of all DSSE residuals at each network MV node. Nodes with frequent occurrences of high residual values should be investigated to determine if these are caused by metering or communication issues, or by incorrect values for network parameters and configuration. Furthermore, all load forecasts are compared against the corresponding recorded real-time values in the post-processing stage. The estimation/forecast errors at each node are used to identify under-performing NN load models, which are then retrained as required. The standard deviations of the MV node load forecasts σ_i are also used as feedback to the DSSE. These values are used to build the measurement weight matrix (4), allowing the DSSE to assign greater importance to load estimates which are known to have better accuracy, improving the overall performance of the DSSE.

CONCLUSIONS

The methodology presented in this paper implements a load estimator and a robust DSSE in a closed-loop configuration. The provision of high-quality load estimates to the DSSE is a difficult problem, due to the inherent variability in MV/LV substation-level measurements. It was demonstrated that non-linear load estimators in this application than linear models for this application. The DSSE was tested using data from an actual MV distribution network and its robustness to measurement noise and errors was demonstrated. It was also shown that the approach can be applied to short-term operational planning in distribution systems. Further work will focus on the testing the DSSE on various types of distribution system (meshed/radial, urban/rural), and on unbalanced MV/LV distribution systems. In addition, further research will be carried out in order to automate the procedures for post-processing of real-time measurement data. This will allow for automatic identification of problems in the input data, and subsequent re-training of load estimation models.

ACKNOWLEDGEMENTS

REFERENCES

[1] D. Haughton, G. Heydt, 2013, "A linear state estimation formulation for smart distribution systems", *IEEE Trans. Power Sys.*, vol. 28 (2), 1187-1195.

[2] J. Wu, Y. He, N. Jenkins, 2013, "A robust state estimator for medium voltage distribution networks," *IEEE Trans. Power Sys.* vol. 28, no. 2, pp. 1008–1016.

[3] S. Grenard, O. Carre, 2012, "Optimal short term operational planning for distribution networks," *CIRED 2012 Workshop*, Lisbon, Portugal, 29-30 May.

[4] B. Hayes et. al., 2014, "Optimal power flow for maximizing network benefits from demand-side management," *IEEE Trans. Power Sys.*, in press.

[5] SmartHG project website, http://smarthg.di.uiniroma.it/

[6] E. Castillo, A. Conejo, R. Pruneda, C. Solares, 2005, "State estimation observability based on the null space of the measurement Jacobian matrix," *IEEE Trans. Power Sys.*, vol. 20 (3), 1656-1658.

[7] R. Singh, B. Pal, R. Jabr, 2009, "Choice of estimator for distribution system state estimation", *IET Gen., Trans., Dist.*, vol. 3 (7), 666-678.

[8] E. Manitsas et. al., 2012, "Distribution system state estimation using an artificial neural network approach for pseudo network modelling," *IEEE Trans. Power Sys.*, vol. 27 (5), 1888-1896.

[9] P. Huber, E. Ronchetti, 2009, *Robust Statistics 2ⁿᵈ Ed.*, Wiley, Hoboken, NJ.

[10] H. Hippert, P. Pedreira, R. Souza, 2001, "Neural networks for short-term load forecasting: a review," *IEEE Trans. Power Sys.*, vol. 16 (1), 44-55.

[11] M. Espinoza et. al., 2007, "Electric load forecasting," *IEEE Control Sys.* vol. 27 (5), 43-57.