IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 21, 2019, accepted January 1, 2020, date of publication January 9, 2020, date of current version January 17, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2965320

Non-Gaussian Residual Based Short Term Load Forecast Adjustment for Distribution Feeders

BRUCE STEPHEN, (Senior Member, IEEE), RORY TELFORD, AND STUART GALLOWAY
Advanced Electrical Systems Research Group, University of Strathclyde, Glasgow G1 1RD, U.K.

Corresponding author: Bruce Stephen (bruce.stephen@strath.ac.uk)

This work was supported by the Engineering and Physical Sciences Research Council Analytical Middleware for Informed Distribution Networks (AMIDiNe) under Grant EP/S030131/1.

ABSTRACT The evolving role for electricity network operators means that load forecasting at the distribution level has become increasingly important, presenting the need for anticipation of the behavior of highly dynamic and diversely distributed loads. The commonly held assumption of Gaussian residuals in forecasting does not always hold for distribution network loads, increasing the uncertainty in balancing a system at this network level. To reduce the operational impact of forecast errors, this paper utilizes different multivariate joint probability distributions to capture the intra-day dependency structure of forecast residuals. Transforming these to the conditional form enables forecast corrections to be made at variable horizons even in the absence of the forecast model. Improvements in accuracy are demonstrated on benchmark load forecast models at distribution level low voltage substations. A practical distribution system application on scheduling embedded energy storage shows substantial reductions in grid imports and hence costs to distribution level customers from utilizing the proposed intraday correction approach.

INDEX TERMS Load modeling, power systems, renewable generation.

I. INTRODUCTION

Distribution networks are becoming increasingly observable through the availability of low cost, low latency monitoring solutions that are required to support the accommodation and management of low carbon technologies, such as distributed energy resources (DER) and electrified heat and transport loads [1]. Failure to accommodate such technologies can result in suboptimal curtailment of generation [2], suboptimal storage scheduling [3], provision of excess spinning reserve higher in the network [4] and evermore costly ancillary services such as short term operating reserve. Addressing this through the adoption of a federated Distribution System Operator (DSO) model will result in loads being balanced further down into the network than is currently the practice; therefore, if substation flexibilities are to be leveraged in such a way, the uncertainties associated with their utilization needs to be reliably anticipated. Loads that possess highly dynamic characteristics can render day ahead forecasts insufficient, and the resultant need for even more localized forecasts means that high computational effort will be required to obtain anticipated demands for a distribution system as a whole [5]. Alterations of forecasts for online optimization of electrical storage charge controllers and characterization of flexibility for demand response participation would be required at the instance errors for previous forecasts were evaluated. The intra-day dependency structure inherent in load profiles offers an element of substructure by which short-term temporal elements of demand can be factored in [6]. A critical part of such modelling is capturing the form of the errors: [7] demonstrated that an assumption of Gaussian residuals was not universal; [8] questioned whether the multi-variate equivalent will hold from a dependency perspective when incorporating multiple forecasts; and [9] noted that intra-day dependency structures can turn out to be sparse.

Unlike national or regional demand, load profiles at the lowest voltage levels of the distribution network have been noted to follow set patterns, even if those patterns don't repeat on a day to day basis [10]–[13] – suggesting intra-day rather than inter-day dependence may be stronger at this level of the network. In [6], it was demonstrated how daily load profiles could be tiled together from recurring sub-profile shapes within the day, highlighting the potential for intra-day substructure to be exploited to decrease demand uncertainty within day. [11] demonstrated that load profile shapes were not typical across distribution networks, even at substation level and identified finite sets of typical behaviors; [13] utilized a similar finding to choose sequences of daily residential load profile shapes from a recurring subset on

The associate editor coordinating the review of this manuscript and approving it for publication was Yang Li.

the basis of learned routine and association with day of week. This assumed a fixed daily load profile shape though. Given strong routine and environmental factors in individual or highly localized energy use, intra-day errors should offer greater potential for improving load profile forecasts in comparison to the inter-day errors on which many forecasting models are based [9]. Co-location of generation and storage behind a common settlement point [14] has led to load profiles taking yet more forms. This results in the probabilistic form of forecast load and its residuals becoming non-Gaussian through skewing or multi-modality.

This paper provides the basis for formal approaches to updating forecasts over a variable time horizon, irrespective of the model that generated them, through capturing the intra-day dependency structures inherent in forecast errors. Where [13] only considered Gaussian loads as marginal in a multi-variate Gaussian load profile, and [11] utilized k-Means, which again had an inherent assumption of Gaussianity, here it is proposed that these assumptions be relaxed in both the form of the dependency, and the form of marginal residual distributions. In particular, this paper proposes the use of copulas to capture joint high dimensional dependence across half hourly forecast errors and adjust day-ahead forecasts according to the corresponding conditional distribution. The method is demonstrated on operational data recorded at 22 separate LV feeders from a total of 8 different 11/0.415kV distribution level secondary substations within an area of the UK. These substations serve a mix of residential, light industrial and commercial premises in both urban and rural conurbations with unreported amounts of embedded generation installed. Monitoring data at this level of the network is not commonly the case with most LV networks in the developed world, thus also demonstrating where this will offer operational benefit without the advantage of performing seasonal de-trending from historical data.

The paper is arranged as follows: Section II describes a distribution feeder test set and outlines potential benchmark load forecast models that have been used for short term load forecasting at higher levels of the network. Section III describes the relation between forecast errors and how copulas can be used adjust day-ahead forecasts; Section IV assesses the proposed copula adjustment methods performance on improving benchmark forecasts for LV feeder data introduced in Section II. A case study in Section V highlights the benefits of the proposed method to reduce grid imports in areas with local embedded generation through improved scheduling battery energy storage (BES). Section VI concludes the paper with observations on the applicability for system operation at LV distribution.

II. DISTRIBUTION LEVEL SHORT TERM LOAD FORECASTING

An increasing challenge with regards to management of distribution networks is ensuring the network operates within relevant constraints while also maximizing output of low carbon energy production. Short term visibility of power flows through accurate forecasts are integral to ensuring these objectives. Short term load forecasting is identified as a forecast which is conducted a day ahead or less [15]. This section summarizes the application of benchmark short-term forecast models to an LV load dataset captured within an area of the UK.

A. LV LOAD DATASET

To highlight the practical challenges of short term forecasts for distribution networks, several contemporary forecasting techniques are demonstrated here on LV demand data. Half-hourly ($T = 48$) peak active power measurements from 22 separate 415V feeders are used for assessing four different forecast models. Feeder data was captured at eight radially connected 11/0.415kV secondary substations in the UK. Figure 1 shows the mean June daily load profiles for each of the 22 feeders.

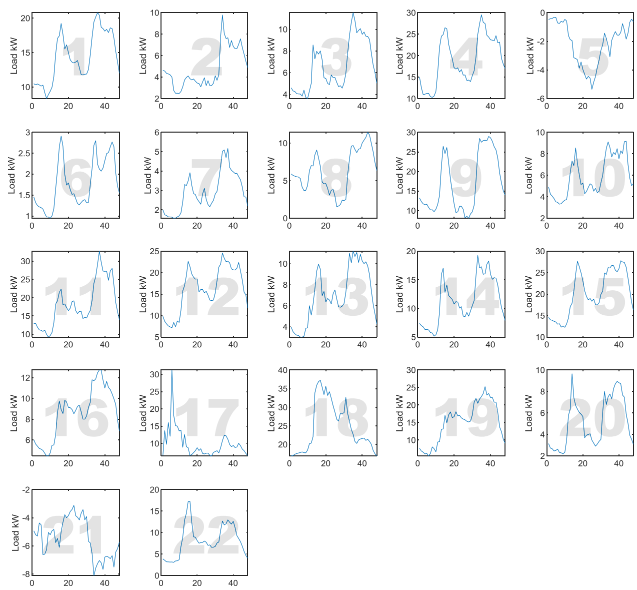


FIGURE 1. Mean daily load profiles for the month of June for the 22 LV feeders considered.

The heterogeneity of these loads is apparent from Figure 1, with a great diversity in both magnitude, shape and time of daily peak load as well as evidence of embedded generation. For the latter point, feeder #5 has the most obvious evidence of embedded generation with a negative demand trough in the middle of the day, while feeder #17 features a high demand peak which abruptly ends with the beginning of daylight hours, suggesting a zero export limit on this feeder. Other feeders such as 9 through to 16 feature more recognizable domestic load profiles albeit at different scales reflecting the extent of load aggregation. These differences motivate the need for a forecast model per feeder.

B. BENCHMARK SHORT-TERM LOAD FORECAST MODELS

The four different benchmark forecast models applied to the LV dataset included: persistence; linear; gradient boost machine (GBM); and, an extreme learning machine (ELM).

The most basic of forecasts is provided from a persistence forecast, which works on the assumption that whatever system behavior is observed at a given time step will also occur in the forecast at the same time interval. To avoid artefacts of weekly routine (for example, weekday/weekend differences and regular scheduled activities) this could be performed using observations from 7 days previous.

Commonly used in recent forecasting works for benchmarking purposes is the 'Vanilla' linear model used by [16]. This model has been utilized at the regional/MV level where long term load and temperature observations are readily available to produce long term seasonal trends. At lower levels of the network [17] notes that temperatures are weak predictors in forecast models. Although this may not be the case in instances where electric space heating is used [18], in the case being considered, domestic gas is available and likely used without exception for space heating, so the temperature terms are dropped. The month variable is also left out owing to the short term nature of this data.

Gradient Boost Machines [19] (GBM) have seen recent application in load forecasting at the regional level where significant benefits in forecasting have been realized [20]. GBM are ensemble methods that address classification and regression tasks using a cascaded approach to combining the outputs of multiple simple models. The 'weak' models are added and learned in stages to progressively refine the output of the model with the shortfall in accuracy of the model being identified through its gradient descent learning procedure.

Regression and classification problems have recently seen treatment by Extreme Learning Machine (ELM) [21]. These use a single layer Neural Network architecture whose hidden layer reduces any function approximation to one of least squares estimation. In [22], day ahead forecasting was undertaken with an ELM.

Contemporary approaches to forecasting including Deep Learning have been avoided in this study owing to their need for large volumes of training data. The selection of base forecast models chosen has been motivated to avoid the functional limitations of a day ahead model from producing just one form of residual [7]. By diversifying the underlying forecast functional approximations, a range of residual substructures is expected to be produced for feeder, season and individual model. This will be used to highlight the level of modeling flexibility required to accurately capture the form of the residuals.

C. BENCHMARK BASE DAY AHEAD FORECAST RESULTS

For each of the 22 LV feeders, the four forecast models described in Section II.B are used for day ahead forecasting using a model learned on the previous month's data. The persistence forecast was the exception to this in that it used the observations from seven days' prior as its forecast. This resulted in $4 \times 22 = 88$ forecasts for 11 separate months i.e. 968 monthly forecasts. From these 968 monthly forecasts, half hourly residuals were calculated, resulting in 46464 half-hourly forecast residuals. Distributions of these residuals for each of the 22 feeders are shown in Figure 2.

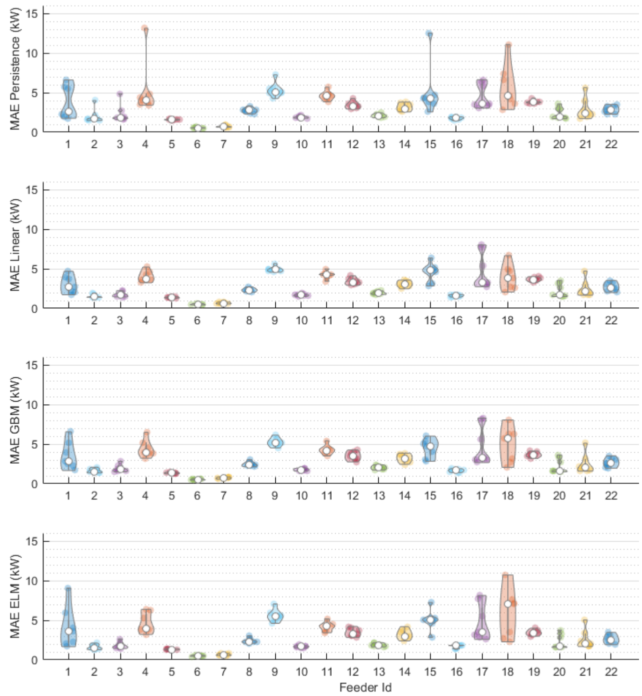


FIGURE 2. Error distributions for four forecast models across 22 feeders from March to December.

Evident from Figure 2 are the very similar errors across the forecast models – median errors are equivalent and the improvements tend to be in the reduction in variance or occurrence of outliers.

III. FORECAST RESIDUAL INTRADAY RELATION

Residuals or the errors resulting from the outturn of a particular electricity load forecast cannot always be assumed to be Gaussian [7]. This is an artefact of both the load and the model used. Identifying intra-day patterns in load and related data [9] can be used to reconcile the inherently inter-day forecast error components that result from models that either account for it linearly or discount it entirely. With the integration of renewable generation at distribution level, the intra-day effects become more complex and result from shorter term phenomena such as routine [10], [23] and meteorological variations. If errors at each forecast time step cannot be assumed to be Gaussian, then their joint intra-day form cannot be assumed to be multi-variate Gaussian and possibly not linear either.

A short term load forecast conducted one day ahead will produce load estimates for T time periods; the residual e of a load l on day d at time t will be:

$$e_{d,t} = \hat{l}_{d,t} - l_{d,t} \tag{1}$$

For any given day the residual will be in

$$e_d \in \Re^T \tag{2}$$

The residual vector for a given day d, would be sampled from a multi-variate distribution of some arbitrary form parameterized by θ:

$$\hat{e}_d \sim f\left(e_d; \theta\right) \tag{3}$$

This would result in a vector of length T which could be added to a subsequent forecast as a correction factor. If f was assumed to be a multi-variate Gaussian distribution with mean μ and covariance $\sum$ then:

$$\hat{e}_{d+1} \sim N\left(\mu, \Sigma\right) \tag{4}$$

This can be used to adjust any given forecast, irrespective of its underlying assumptions, as follows

$$\tilde{l}_{d+1} = \hat{l}_{d+1} - \hat{e}_{d+1} \tag{5}$$

A further convenience of the use of a joint Gaussian distributed residual vector is that as errors are incrementally observed, the conditional distribution, also Gaussian in its form [20], can be used to make intermediate residual predictions [24]; the mean and variance of this predictive distribution will be as follows:

$$\bar{\mu} = \mu_t + \Sigma_{\mathrm{AB}}\Sigma_{BB}^{-1}\left(e_{d,t-h} - \mu_{t-h}\right) \tag{6}$$

$$\bar{\Sigma} = \Sigma_{AA} - \Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{BA} \tag{7}$$

where A is the subset of time period(s) within the day between time t and horizon h forming the conditional distribution (the forecast) and B is the set of time period(s) $1, 2, \ldots, t-1$ being conditioned on (the observation). An error forecast can then be obtained either from the predictive mean (6) as in [24] or as shown in [8] by sampling from:

$$\hat{e}_{d,t+h} \sim N\left(\bar{\mu}, \bar{\Sigma}\right) \tag{8}$$

And an adjustment to the forecast as in (5) can be carried out as the errors for each conditioning advance are observed. The intuitive partitioning of the covariance for prediction in (6) and (7) is illustrated in Figure 3 where the pairwise intra-day Gaussian covariances are shown as cells in T by T matrices. The scale of Figure 3 ranges from strong positive covariance (red) through to strong negative covariance (blue). Positive covariance denotes mutual rises in errors at pairs of times – generally, these are consecutive times and capture over or underestimated extreme in demand. Negative covariance denotes an error rising or falling contrary to an error at another time period. This can be attributed to demand extremes shifting in time with the displacement being caught by the forecast as an opposing error at the next timestep. Through the tessellations in the covariance matrices, Figure 3 highlights that some intra-day errors will have no bearing on those observed later in the day. This is based on a key, and limiting, assumption that residuals are Gaussian distributed and they are linearly related – this assumption will obscure relations that exist in a different, non-Gaussian, form. In [7], the basis for errors or residuals of forecasts being assumed to be Gaussian distributed was tested. Evaluating the Gaussianity of the marginal distribution of the residuals will indicate whether a joint Gaussian assumption can be made across all T intervals in the forecast. The next two sections, carry out analysis on the residuals to understand if the Gaussian assumption holds for all feeders and all forecast models.

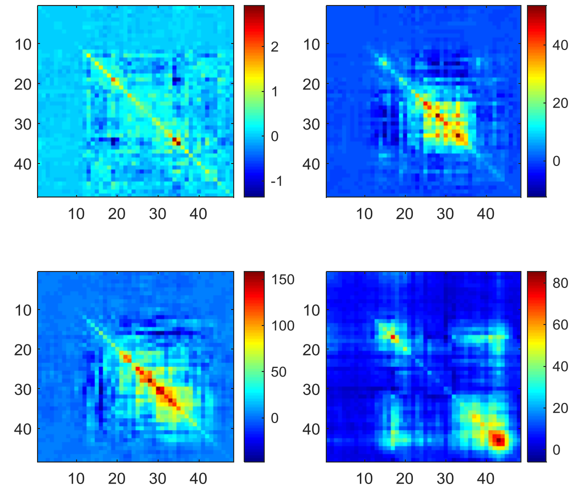


FIGURE 3. Intra-day covariance matrix from joint distribution of error residuals from persistence load forecasts for feeders 7 (top left), 8 (top right), 9 (bottom left) and 18 (bottom right). Some pairs of time periods are highly correlated on some feeders (9 and 18), others have weak short term interactions (8) and others have very little linear dependency structure evident (7). T is 48 for this day ahead forecast.

A. MARGINAL DISTRIBUTION OF RESIDUALS

For each forecast described in Section II, the Gaussianity of monthly residuals is tested for each feeder. The Lillefors test [25] is used to determine if the Null Hypothesis that the data is Gaussian should be rejected with a significance greater than 5%. Figure 4 shows both the empirical distribution of residuals and their cumulative distribution functions for a

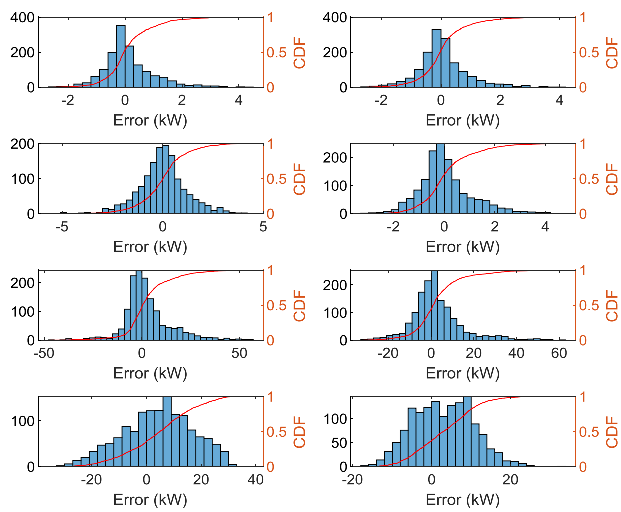


FIGURE 4. 8 forecast residual distributions for June Linear (left) and ELM (right) day ahead forecasts for feeders (top to bottom) 1, 2, 3 and 4. All distributions are deemed non-Gaussian by their failure of the Lillefors test.

selection of 8 forecasts. Figure 4 shows the entire selection of these residuals are evidently non-Gaussian, either exhibiting skewness, multi-modality or high kurtosis that negates their Gaussianity. This is in alignment with the findings of [7] and [13]. In response, [7] proposed partitioning residuals by hour to make them more Gaussian. Following this strategy, the same approach is taken with the 24 separate hourly residual distributions per forecast per feeder and is presented in Table 1. Table 1 demonstrates that this partitioning improves matters, with the highest number of Gaussian hourly residuals being 86% - this was generated from the Persistence forecast in July. An alternative would be to consider either the empirical distribution or a Kernel Density Estimate of the marginal distribution of load. This approach would accommodate non-stationary effects exhibited through the hourly residuals as skewed or kurtotic distributions.

TABLE 1. % of non-gaussian hourly load forecast residuals.

	Persistence	Linear	GBM	ELM
March	27	36	39	38
April	22	25	29	28
May	22	25	21	22
June	17	21	25	26
July	14	22	17	21
August	19	27	28	29
September	21	29	28	27
October	35	41	46	41
November	19	24	25	31
December	20	26	23	24

B. JOINT DISTRIBUTION OF INTRA-DAY RESIDUALS

If marginal distributions cannot be considered Gaussian, as Table 1 suggests, then this rules out a resulting joint distribution across a day from being Gaussian. This section considers intra-day dependency structure for both multi-variate distributed residuals and their equivalent copula based joint distribution. Joint distribution of residuals were examined by [8], who also highlighted a more complex dependency structure existed than was afforded by a multi-variate Gaussian.

The use of copulas has become popular in situations where a multi-variate distribution of an appropriate form is not available that can accommodate either the distribution or dependency relation of its constituent variables. In probability theory a copula, C, enables the decomposition of the joint distribution function into its marginal distributions and a dependency structure [27]. Although the copula relates uniformly distributed variables, arbitrarily distributed variables can be transformed into uniform ones through their own cumulative density functions. Therefore, a cumulative density function can be assimilated using C as:

$$F_{X1X2}(x_1, x_2) = C(u_1, u_2) = C\left(F_{x_1}(x_1), F_{x_2}(x_2)\right) \quad (9)$$

u_1 and u_2 will then be uniformly distributed in [0,1], irrespective of the original marginal distributions of x and y. The joint probability density function f is obtained:

$$f_{X_1X_2}(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot C(u_1, u_2) \quad (10)$$

As any random variable can be converted into a uniformly distributed one by passing it through its cumulative density function, the joint distribution of a T dimensional residual could be represented via a copula as:

$$f(e_d; \theta) = C\left(F_1(e_1), \ldots, F_T(e_T)\right) \prod_{t=1}^{T} f(e_t) \quad (11)$$

The Gaussian copula is given by:

$$C\left(u_{d,1} \ldots u_{d,T}; R\right) = \Phi\left(\phi^{-1}\left(u_{d,1}\right), \ldots, \phi^{-1}\left(u_{d,T}\right); R\right) \quad (12)$$

Φ is the multi-variate Normal CDF and u is a uniformly distributed variable obtained by passing any variable, such as a forecast error, through its assumed cumulative density function F:

$$u_{d,t} = F_t\left(e_{d,t}\right) \quad (13)$$

In coupling intra-day residuals, the Gaussian copula (12) would be parameterized by a T by T covariance matrix R that captures pairwise intra-day dependency in its off diagonal elements. R can be estimated as the sample covariance after the inverse standard Normal cumulative density function is used to transform the uniform variable into a one:

$$y_{d,t} = \phi^{-1}\left(u_{d,t}\right) \quad (14)$$

The significant advantage here is that the assumption of marginally Gaussian distributed errors at all times of the day is relaxed. Given the results outlined in Table 1, the use of copulas to capture intraday dependence, and thus adjust short-term forecasts may offer significant benefits.

IV. FORECAST PERFORMANCE: DAY AHEAD AND INTRA-DAY CORRECTED

This section describes results of applying both a Gaussian and copula based error joint distribution as a means of adjusting forecasts throughout a day. This process begins with the usual parameter estimation for the day ahead forecast model (see Figure 5) irrespective of the forecast model used.

For demonstration purposes, one month of data is used to learn the forecast model parameters for each feeder; a set of forecasts are then produced in-sample from this month in order to obtain expected values for forecast errors at each time of day – 48 half hours as used in this paper. Each model for each feeder is then used to make day ahead forecasts for the subsequent month. The residuals obtained in sample are used to fit marginal distributions for the half hourly errors (here, this includes both Gaussian for the benchmark and Kernel Density Estimates for the arbitrary distribution form case) as well as estimate the intra-day dependency structure between

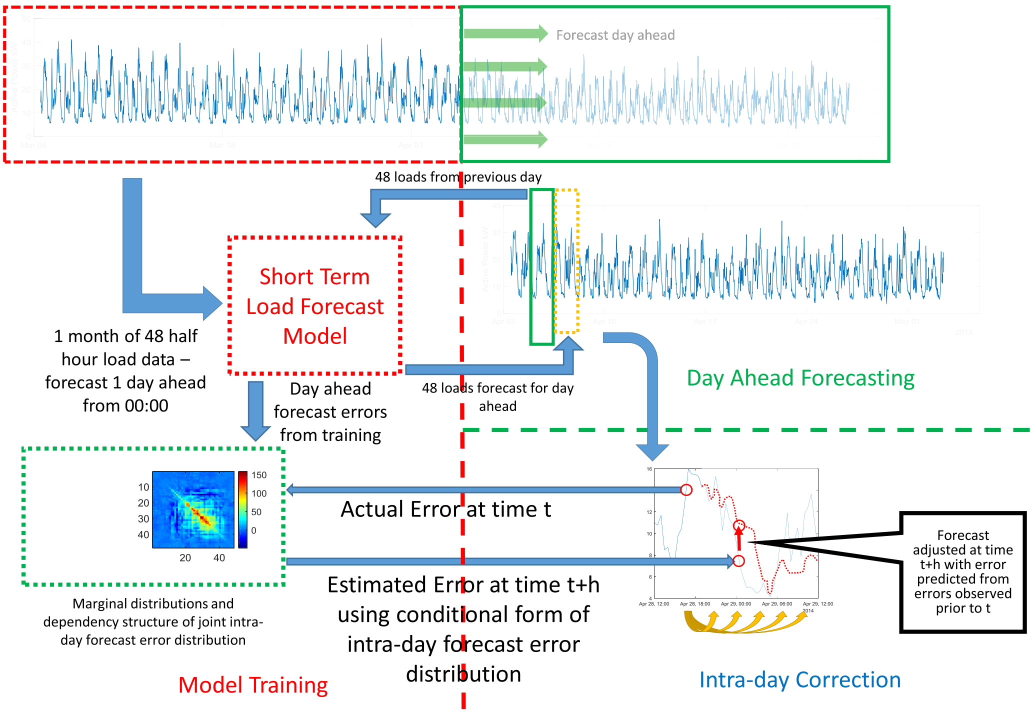


FIGURE 5. Procedure for learning forecast model, fitting the residual distribution and making intraday corrections to day ahead load forecasts.

them. Again, for the purposes of this paper, this is a comparison between a multivariate Gaussian and a copula with marginal distributions of an arbitrary form. The right-most part of Figure 5 shows how these two models may be used operationally, with the short term load forecast being applied first at midnight to generate a day's worth of predicted loads based on the previous days loads. Secondly, the conditional form of the intra-day error joint distribution can be used to predict residuals at later times in the day from those observed earlier, and adjust the forecast accordingly.

For the 88 monthly day-ahead forecasts described in Section II, the Mean Absolute Error (MAE) is calculated at each half hourly demand out-turn and used to correct the day-ahead forecast at horizons of between 1 half hour and 12 half hours using only the intra-day error joint distributions. The joint distributions are estimated from the day ahead residuals of the training data taken from the previous month, yielding a mean vector and covariance matrix for the multivariate Gaussian estimate of errors, and a covariance matrix R for the Gaussian copula based error distribution. Using (6), the conditional form of these distributions is used as a predictor for errors within the day, with the set of predictions given by:

$$A = \{t\} \tag{15}$$

and the set of predictors given by:

$$B = \{t - h\} \tag{16}$$

An intra-day correction over horizon h can then be made with the following predictor assuming Gaussian residuals:

$$\hat{e}_{d,t} = \mu_t + \Sigma_{\mathrm{AB}} \Sigma_{BB}^{-1} \left(e_{d,t-h} - \mu_{t-h}\right) \tag{17}$$

While the following is the basis for the equivalent copula based error prediction:

$$\hat{e}_{d,t} = F_t^{-1} \left(R_{\mathrm{AB}} R_{BB}^{-1} \phi^{-1} \left(F_{t-h}\left(e_{d,t-h}\right)\right)\right) \tag{18}$$

The estimates from (17) and (18) are then used in (5) to correct the associated day ahead forecast within the day. Figure 6 illustrates an example of how this proposed methodology improves base forecasts. In particular, it highlights how MAEs differ between a base ELM forecast and the adjusted forecasts across the 22 LV feeders throughout the month of June, Within Figure 6, h is the adjusted horizon and ranges between 1 to 12 half hour intervals.

From Figure 6, the reduction in MAE identifies that longer period intra-day dependency is captured better with the copula with kernel density marginals: at worst, the error follows that of the Gaussian correction until the 3-5 hour horizon and in many cases betters it. The larger reduction in errors for the copula based model is indicative that both non-stationary behavior and non-linear dependence are present and accounted for in corrected forecasts. Figure 7 shows the benefit of forecast corrections in terms of forecast Mean Absolute Percentage Error (MAPE).

For Figure 7a the long term benefit of any forecast reduction over the base forecast is clear, although using the correct residual assumptions embodied by the copula offers the greater improvement. Figure 7b shows that during the summer months, a Gaussian intra-day correction actually makes the forecast worse in some cases and in the best case only offers a slight reduction. This can be attributed to greater embedded PV generation during May through to

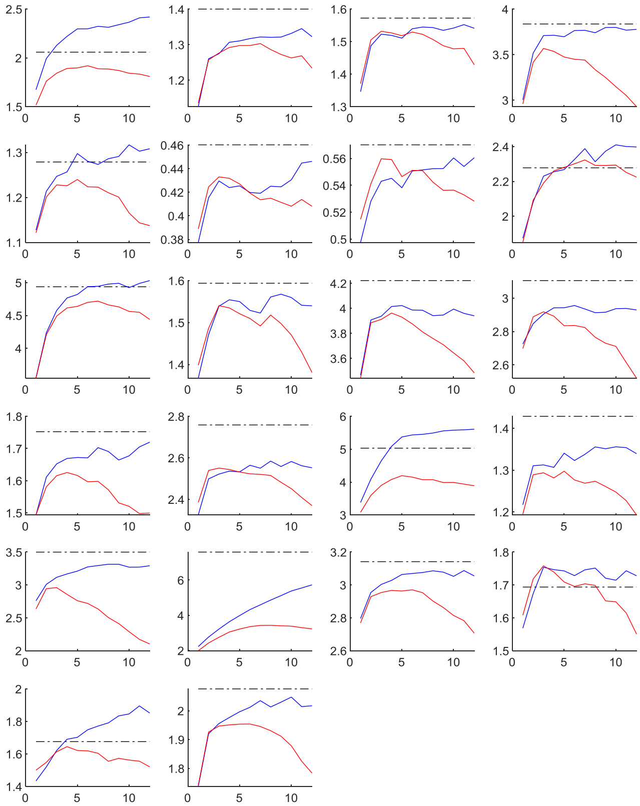


FIGURE 6. MAE (y-axis, in kW) after load residual correction (x-axis, 30 minute intervals) for 22 June ELM forecasts: black line is the uncorrected forecast error, blue line is the Gaussian corrected forecast error for 1-12 half hour look ahead, red line is the Gaussian copula corrected forecast error for 1-12 half hour look ahead.

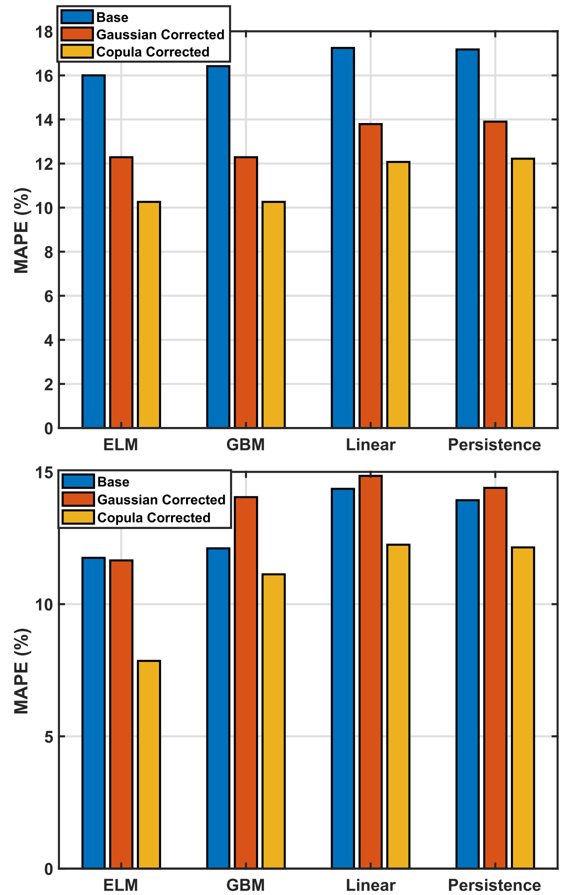


FIGURE 7. (a- top) Reductions in forecast Mean Absolute Percentage Error for 11 months of operation and (b- lower) for just months May-August.

August which could result in a bi-modal residual distribution which the Gaussian approximation cannot handle. The resulting incorrect error distribution then provides a poor forecast adjustment. Overall, these preliminary results show the potential benefit for reduction in forecast uncertainty of the proposed copula based adjustment, but additional factors need to be considered in order to quantify operational value beyond forecast error reduction.

V. INTRA-DAY CORRECTED SCHEDULING OF BATTERY ENERGY STORAGE: A CASE STUDY

A case study on BES scheduling at LV levels is used to further demonstrate operational benefits and practical application of the intra-day forecast adjustment method proposed in this paper. The case study analyses the benefits that an improved load forecast will have in terms of dispatching energy storage in LV networks with embedded generation.

A. LV SYSTEM MODEL

The LV dataset, previously described in Section II, consisting of half-hourly active power load measurements captured at 22 feeders from 8 different secondary substations within the UK was used to develop an LV system model. The 8 different substations are radially connected to an 11kV primary feeder,

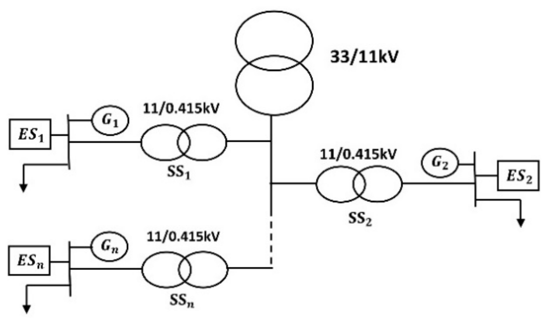


FIGURE 8. One-line diagram of LV network model used for local import minimization case study.

and each secondary substation has a variable number of 415V feeders distributing power to homes and businesses – the number of feeders at each secondary substation range from a maximum of 5 to a minimum of 1. The one-line diagram in Figure 8 illustrates the basic radial architecture of the LV network where separate 415V feeders at each secondary substation are represented as a lumped load and $n = 8$ is the total number of radially connected substations. Separate feeder IDs at each substation outlined in Table 2 refer to the feeder numbers shown in Figure 1.

PV embedded generation and BES was also modelled at each secondary substation, as shown in Figure 6. PV generation was modelled at each secondary substation location using global horizontal irradiance (GHI) data [28] and a PV simulation toolset [29]. Model parameters are summarized in Table 2.

TABLE 2. LV network model parameters.

ES	ES_1	ES_2	ES_3	ES_4	ES_5	ES_6	ES_7	ES_8
Capacity (kWh)	280	200	140	120	60	60	50	140
PV	G_1	G_2	G_3	G_4	G_5	G_6	G_7	G_8
Rating (kW)	140	100	70	60	30	30	25	70
Feeder	SS_1	SS_2	SS_3	SS_4	SS_5	SS_6	SS_7	SS
IDs	10, 11, 12, 13, 14	16, 17, 18, 19	6,7,8, 9	2,3,4, 5	20, 21	1	15	22

B. BATTERY ENERGY STORAGE SCHEDULING

The load forecasts described in Section IV, were combined with a basic persistence PV generation forecast to schedule BES dispatch. The BES scheduling policy objective was based on the method proposed in [30], and aims to minimize both energy import to the local load and energy export from the local generation. Ideally, under this objective, BES should charge at the points in time where generation exceeds local demand, and discharge appropriately when demand exceeds generation. Under this objective, the role of accurate forecasts is vital in dispatch scheduling, particularly in instances where embedded generation output and demand are closely aligned and an inaccurate forecast will result in an incorrect BES dispatch signal. Reducing uncertainty with regards to the system dispatch, and hence flexibility, will become increasingly critical at LV levels as distribution network operators transition to system operators - this case study highlights how the proposed adjusted Gaussian copula load forecasting method can benefit overall system management in a BES scheduling context.

In total, 12 load forecasts – the four base line forecasts described in Section II, plus their Gaussian error adjusted and Gaussian copula adjusted equivalents, described in Section IV - were used with a persistence PV generation forecast to produce 12 different ES dispatch schedules for each battery. In the instances where an adjusted load forecast is used, the basic dispatch schedules are updated at each half hour advance throughout the day using receding horizon control; for the base forecasts, schedules are calculated day ahead.

Dispatch schedules for each of the 8 batteries in Figure 6 assumed charge and discharge limits were 90% and 20% times their maximum storage capacity respectively – ES capacities, outlined in Table 2, were sized to double the PV installed capacity to ensure that battery capacity limitations would not significantly impact the ability to store the local generation. Note that this analysis does not consider the optimization of ES capacity. All batteries were modelled with a round-trip efficiency of 75%. The 12 different ES schedules were produced across the three summer months (June, July and August) in the UK when GHI peaks.

C. ANALYSIS OF ES DISPATCH SCHEDULING

The 12 ES dispatch schedules for each battery were assessed in terms of energy imports at each of the 8 substations throughout the three-month period. Figure 9 illustrates import reductions across all secondary substations between the base forecasts and the Gaussian error adjusted forecasts, and the import reductions between the base forecasts and the proposed Gaussian copula adjusted forecasts.

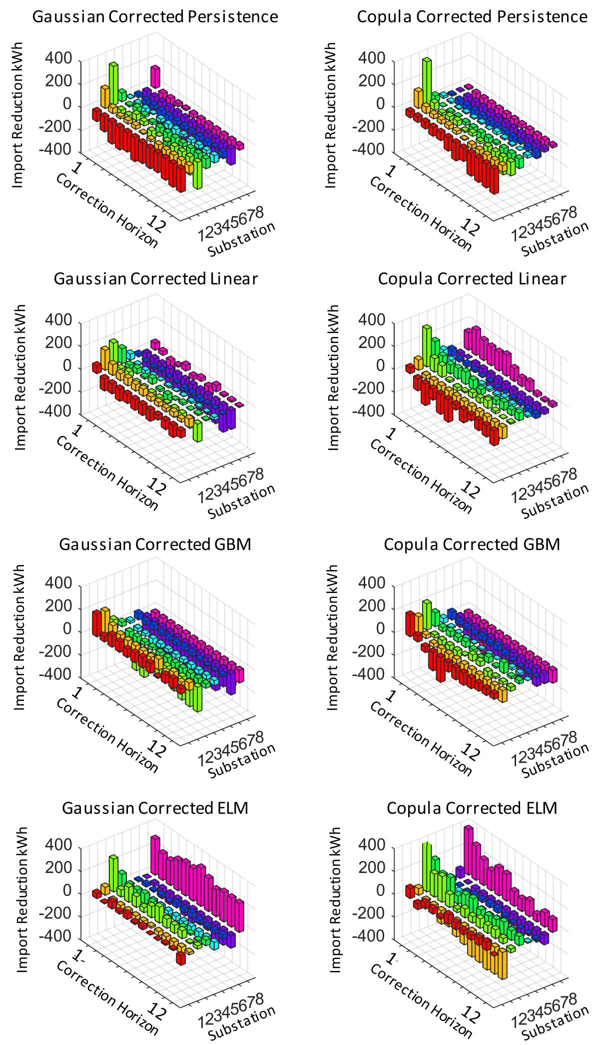


FIGURE 9. (a- left) Reductions in energy imports for each substation across three months when using a Gaussian error adjusted load forecast for ES dispatch scheduling as opposed to a base day-ahead forecast and (b- right) import reductions when using a Gaussian copula adjusted load forecast for ES dispatch as opposed to a base forecast.

Figure 8 quantifies the reductions in energy imports using corrected forecasts highlighting it is particularly effective across short horizons. Figure 9 also demonstrates that although a forecast correction may have been possible, the other uncertainties in the system (e.g. generation forecasts, naive storage scheduling) have resulted in no or increased imports over the base forecast. Correcting an ELM forecast using the proposed copula based method with a correction horizon of 1 half hour would yield total import reductions across the 3-month period for all 8 substations of 1.385MWh – in comparison, adjusting with a Gaussian model would result in reductions of just 0.845 MWh. Similarly, adjusting a persistence forecast would result in a reduction of 0.515MWh across all eight substations while adjusting using the Gaussian model would result in only 0.476MWh. Gaussian correction still offers reduction potential but it is sub-optimal.

Figure 10 highlights the periods where both adjusted forecasts either increase (negative reduction) or reduce (positive reduction) energy imports across substations in comparison to base forecasts. Within Figure 10, white cells indicate where imports are reduced and black cells indicate where imports are increased. Improvements over base are minimal for the persistence forecast with the Gaussian adjustment only improving 15% of forecasts for any horizon up to 12 hours. The copula approach is marginally better at 25% over the base forecast but manages 71% better import reductions over the Gaussian. The results highlight that, when adjusting a persistence forecast, any reduction in import levels would be confined to cases where a short adjustment horizon is adopted. Comparatively, the copula based adjustment method reduces imports across the 8 substations particularly well when utilizing a linear and ELM forecasts. Overall, as shown in Figure 10, the copula adjusted method reduces imports to a greater extent than a Gaussian based method, implying non-Gaussian residuals.

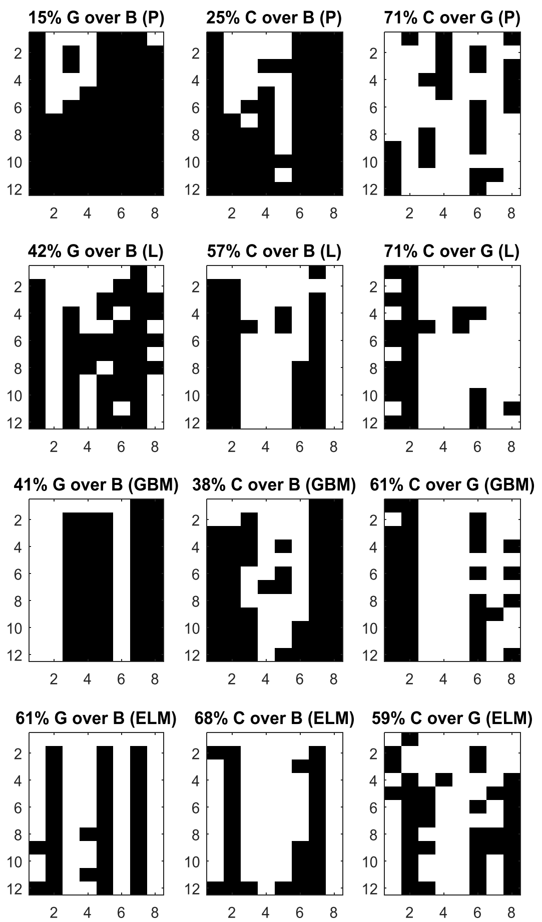


FIGURE 10. Adjustment improvements for G(aussian) over B(ase), C(opula) over B(ase) and C(opula) over G(aussian) for P(ersistence), L(inear), GBM and ELM day ahead forecasts. Columns represent substations while rows represent forecast horizon adjustment in terms of half hour interval. White cells indicate an improvement for the particular adjustment horizon for a particular substation.

As mentioned previously, several uncertainties on top of accurate load forecasting can affect how successful a particular BES schedule is in meeting specified objectives. Such challenges are compounded by uncertainty in demand at LV levels. This case study has served to show how improving forecasts of load at LV can take steps towards reducing uncertainty and enhancing operational benefits.

D. DSO PERSPECTIVE

As distribution networks feature ever more generation and potentially, storage penetrations, the need for balancing and ancillary service procurement deep into the network increases [31]. Owing to the number of LV feeders utility companies own, any forecasting method will need to be optimized for speed and memory use in order to scale up for deployment across distribution networks. Understanding each resulting forecast model capability in terms of accuracy and acting on this will be essential from a practical viewpoint. A cautious system operator may then choose to maintain several update horizons to monitor emerging flexibilities in the form of turn up or turn down; these in turn may align with the behavior of a collection of feeders operating as a virtual balancing unit. In practical terms, forecasts may arrive as a service and the underlying limitations may not be apparent at the point of receipt. Being able to take deterministic or point forecasts and develop a probabilistic residual framework around them would be a useful proposition, particularly if forecasts were being received from multiple providers. While not going as far as reverse engineering the forecast, the intra-day boost as proposed here, would place both the error bars and an ultra-short term forecasting capability over and above the base forecast. Resulting biases stemming from monthly forecast differences and limited training data (which is always likely to be a problem for distribution system operators in their initial years) could be cancelled out by the estimates of the residuals.

VI. CONCLUSION

Distribution system operation presents a number of challenges to load forecasting when considering the number and variability of distribution feeders. This paper has investigated the actual form of forecast errors on the LV side of distribution networks and looked at how making more realistic assumptions regarding their distribution and intra-day dependency can be represented using copulas. Modelling this as a forecast agnostic multivariate error distribution yields a conditional distribution from which predictions of error within the day can be made at arbitrary horizons. The resulting predictive model has improved accuracy with minimal data requirements for day ahead forecasts. The practical benefit of this has been shown to be the accommodation of intra-day variability in highly a localized system context leading to increased utilization of renewable energy to meet local demand. As balancing services are increasingly required at lower voltage levels of distribution networks, this will be a commercial and operational benefit if implemented. Production grade implementation will require computational efficiency to be considered which will necessitate (13) to be optimized in some manner. The Kernel Density Estimates used to estimate the true form of the residual distributions are computationally demanding but previous work has shown how they can be made to run significantly faster [32]. In terms of future directions for intra-day modelling of residuals, a key challenge is the dimensionality invoked by the resolution of the model. Copula models that can work with high dimensional data include Vines [33]: Vines support factorization of joint densities and copula theory to provide flexible dependency structures in high dimensions. In [34], it was demonstrated how load profile non-stationarity and complex dependency could be accommodated in high dimensions, such an approach applied to residuals could potentially be applied to forecast residuals to attain higher levels of forecast error reduction at longer time horizons.

REFERENCES

[1] A. D. Peacock, E. H. Owens, S. Roaf, D. W. Corne, M. Dissanayake, P. G. Tuohy, B. Stephen, and S. Galloway, "Autarkic energy systems: Balancing supply and demand with energy storage and controls in local energy micro-grids," in *Proc. Asia–Pacific Sol. Res. Conf.*, Sydney, NSW, Australia, Aug. 2014.

[2] T. Stetz, F. Marten, and M. Braun, "Improved low voltage grid-integration of photovoltaic systems in Germany," *IEEE Trans. Sustain. Energy*, vol. 4, no. 2, pp. 534–542, Apr. 2013.

[3] N. Gast, D.-C. Tomozei, and J.-Y. Le Boudec, "Optimal generation and storage scheduling in the presence of renewable forecast uncertainties," *IEEE Trans. Smart Grid*, vol. 5, no. 3, pp. 1328–1339, May 2014.

[4] J. W. Taylor and M. B. Roberts, "Forecasting frequency-corrected electricity demand to support frequency control," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 1925–1932, May 2016.

[5] M. Rowe, T. Yunusov, S. Haben, W. Holderbaum, and B. Potter, "The real-time optimisation of DNO owned storage devices on the LV network for peak reduction," *Energies*, vol. 7, no. 6, pp. 3537–3560, May 2014.

[6] M. Chaouch, "Clustering-based improvement of nonparametric functional time series forecasting: Application to intra-day household-level load curves," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 411–419, Jan. 2014.

[7] J. Xie, T. Hong, T. Laing, and C. Kang, "On normality assumption in residual simulation for probabilistic load forecasting," *IEEE Trans. Smart Grid*, vol. 8, no. 3, pp. 1046–1053, May 2017.

[8] A. Müller, A. Lenkoski, and T. L. Thorarinsdottir, "Multivariate probabilistic forecasting using ensemble Bayesian model averaging and copulas," *Quart. J. Roy. Meteorol. Soc.*, vol. 139, no. 673, pp. 982–991, Apr. 2013.

[9] F. Ziel, "Forecasting electricity spot prices using lasso: On capturing the autoregressive intraday structure," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4977–4987, Nov. 2016.

[10] J. Kwac, J. Flora, and R. Rajagopal, "Household energy consumption segmentation using hourly data," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 420–430, Jan. 2014.

[11] R. Li, C. Gu, F. Li, G. Shaddick and M. Dale, "Development of low voltage network templates—Part I: Substation clustering and classification," *IEEE Trans. Power Syst.*, vol. 30, no. 6, pp. 3036–3044, Dec. 2015.

[12] A. Mutanen, M. Ruska, S. Repo and P. Jarventausta, "Customer classification and load profiling method for distribution systems," *IEEE Trans. Power Del.*, vol. 26, no. 3, pp. 1755–1763, May 2011.

[13] B. Stephen, X. Tang, P. R. Harvey, S. Galloway, and K. I. Jennett, "Incorporating practice theory in sub-profile models for short term aggregated residential load forecasting," *IEEE Trans. Smart Grid*, vol. 8, no. 4, pp. 1591–1598, Jul. 2017.

[14] J. Hardy, "Non-traditional business models: Supporting transformative change in the energy market," Office Gas Energy Market, London, U.K., Tech. Rep., Feb. 2015. [Online]. Available: https://www.ofgem.gov.uk/ofgem-publications/93586/non-traditionalbusinessmodelsdiscussion paperpdf

[15] X. Sun, P. B. Luh, K. W. Cheung, W. Guan, L. D. Michel, S. S. Venkata, and M. T. Miller, "An efficient approach to short-term load forecasting at the distribution level," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2526–2537, Jul. 2016.

[16] T. Hong, P. Wang, and H. L. Willis, "A Naïve multiple linear regression benchmark for short term load forecasting," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2011, pp. 1–6.

[17] S. Haben, G. Giasemidis, F. Ziel, and S. Arora, "Short term load forecasting and the effect of temperature at the low voltage level," *Int. J. Forecasting*, vol. 35, no. 4, pp. 1469–1484, Oct. 2019.

[18] S. Gill, M. Dolan, D. Frame, and G. G. Ault, "The role of the electric heating and district heating networks in the integration of wind energy to Island networks," *Int. J. Distrib. Energy Resour.*, vol. 7, no. 3, pp. 245–263, 2011. ISSN 1614-7138.

[19] J. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Stat.*, vol. 25, no. 5, pp. 1189–1232, 2001.

[20] J. R. Lloyd, "GEFCom2012 hierarchical load forecasting: Gradient boosting machines and Gaussian processes," *Int. J. Forecasting*, vol. 30, no. 2, pp. 369–374, Apr. 2014.

[21] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012.

[22] S. Li, L. Goel, and P. Wang, "An ensemble approach for short-term load forecasting by extreme learning machine," *Appl. Energy*, vol. 170, pp. 22–29, May 2016.

[23] S. Higginson, E. McKenna, and M. Thomson, "Can practice make perfect (models)? Incorporating social practice theory into quantitative energy demand models," in *Proc. 3rd Behave Energy Conf. Paradigm Shift Energy Efficient Energy Reduct. Soc. Change*, Oxford, U.K., Sep. 2014, pp. 3–4 and 17.

[24] E. Käärik and M. Käärik, "Modeling dropouts by conditional distribution, a copula-based approach," *J. Stat. Planning Inference*, vol. 139, no. 11, pp. 3830–3835, Nov. 2009.

[25] H. Abdi and P. Molin, "Lilliefors/Van Soest's test of normality," in *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA, USA: Sage, 2007, pp. 540–544.

[26] E. Carpaneto, G. Chicco, R. Napoli, and M. Scutariu, "Electricity customer classification using frequency-domain load pattern data," *Int. J. Electr. Power Energy Syst.*, vol. 28, no. 1, pp. 13–20, Jan. 2006.

[27] P. Embrechts, A. McNeil, and D. Straumann, "Correlation and dependence in risk management: Properties and pitfalls," in *Risk Management: Value at Risk and Beyond*. Cambridge, U.K.: Cambridge Univ. Press, 2002.

[28] *Modern-Era Retrospective Analysis for Research and Applications, Version 2*, North Amer. Space Agency, Washington, DC, USA. Accessed: Nov. 16, 2018.

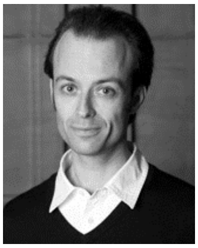
[29] Sandia National Lab. *PV Toolbox*. Accessed: Nov. 16, 2018. [Online]. Available: https://pvpmc.sandia.gov/applications/pv_lib-toolbox/

[30] H.-I. Su and A. E. Gamal, "Modeling and analysis of the role of energy storage for renewable integration: Power balancing," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4109–4117, Nov. 2013.

[31] J. Silva, J. Sumaili, R. J. Bessa, L. Seca, M. Matos, and V. Miranda, "The challenges of estimating the impact of distributed energy resources flexibility on the TSO/DSO boundary node operating points," *Comput. Oper. Res.*, vol. 96, pp. 294–304, Aug. 2018.

[32] V. C. Raykar, R. Duraiswami, and L. H. Zhao, "Fast computation of kernel estimators," *J. Comput. Graph. Statist.*, vol. 19, no. 1, pp. 205–220, Jan. 2010.

[33] T. Bedford and R. M. Cooke, "Probability density decomposition for conditionally dependent random variables modelled by vines," *Ann. Math. Artif. Intell.*, vol. 32., nos. 1–4, pp. 245–268, 2001.
[34] M. Sun, I. Konstantelos, and G. Strbac, "C-vine copula mixture model for clustering of residential electrical load pattern data," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 2382–2393, May 2017.



BRUCE STEPHEN (Senior Member, IEEE) received the B.Sc. degree in aeronautical engineering from the University of Glasgow, U.K., in 1997, the M.Sc. degree in computer science from the University of Strathclyde, U.K., in 1998, and the Ph.D. degree from the University of Strathclyde, in 2005. He is currently a Senior Research Fellow with the Institute for Energy and Environment, University of Strathclyde. His research interests include power system condition monitoring, renewable integration, characterizing low voltage network behavior, and characterization of demand in electrical distribution networks. He is a Chartered Engineer and a Fellow of the Higher Education Academy.

RORY TELFORD received the B.Eng., M.Sc., and Ph.D. degrees in electronic and electrical engineering from the University of Strathclyde, Glasgow, U.K., in 2008, 2011, and 2017, respectively. He is currently a Research Associate with the Institute for Energy and Environment, University of Strathclyde. His research interests include application of AI techniques, data-driven fault diagnostics, and power system modeling and analysis.

STUART GALLOWAY received the M.Sc. and Ph.D. degrees in mathematics from the University of Edinburgh, in 1994 and 1998, respectively. He is currently a Professor with the Institute for Energy and Environment, University of Strathclyde. His industry research work on compact energy systems has been applied to several to several domains across land, sea, and air. His research interests include power system optimization, numerical methods, and the simulation of novel electrical architectures.