Modelling the Interdependence of Multiple Electricity Markets in the Distribution System Aggregator Bidding

Mohammad Afkousi-Paqaleh, *Graduate Student Member, IEEE*, Mohammad Jafarian, and Andrew Keane, *Senior Member, IEEE*

***Abstract*—The volatility of market prices and the interdependence of multiple markets make it challenging for Distribution System Aggregators (DSAs) to model these prices. In this paper, a novel statistical model based on Gaussian process regression and mutual information screening technique is developed. This model is able to predict different market prices and quantify their uncertainty whilst incorporating the interdependence of different markets. The proposed model is employed to assist DSAs with market price modelling. Price scenarios for various markets generated by this model make it viable to formulate the optimal involvement of DSAs in multiple markets as a stochastic multi-step two-stage problem. Other than providing a set of scenarios that efficaciously model multiple electricity market prices, after the clearing of each market, the proposed model leverages market clearing results to improve the accuracy of price prediction of subsequent markets. Extensive simulation results on large price datasets demonstrate that the proposed methodology will result in a considerable increase in the profit of the DSA compared to state-of-the-art price prediction approaches.**

***Index Terms*—Bidding strategy, distribution system aggregator, electricity markets, Gaussian process regression, uncertainty modelling.**

I. Introduction

As MORE Distributed Energy Resources (DERs) are installed in the distribution network and with the ongoing electrification of heating and transport sectors, the distribution system is increasingly encountered with more flexibility and control potentials. An underappreciated opportunity for the power systems pursuing to drive the energy transition lies in the identification of these potentials. These potentials could be aggregated by a Distribution System Aggregator (DSA) and offered to the electricity markets. DSAs should determine an optimal bidding strategy for participation in multiple energy markets to maximise their profit. Nevertheless, determining an optimal bidding strategy is difficult due to the need to account for the uncertainty of market prices and resources, the interdependence of different markets, and the network's operational constraints.

A. Literature Review

To tackle the bidding strategy challenges of an aggregator of small prosumers, [1] applies a two-stage stochastic optimisation technique to minimise the net cost of buying and selling energy. In [2], a model of the Day-Ahead (DA) bidding strategy for a load aggregator is presented, wherein Demand Response (DR) is utilised to mitigate the risk. A new distributionally robust optimisation is proposed in [3] to design a collaborative DA bidding strategy for a DSA of intermittent resources. It is shown that the collaborative bidding strategy yields a more significant financial gain than independent bidding. The authors in [4] propose a secure and effective bidding strategy and Real-Time (RT) operation of a DSA that incorporates distribution network constraints. However, the above studies, [1], [2], [3], [4], have failed to address the uncertainties of market prices.

A growing body of literature recognises the importance of factoring in the price uncertainties in the bidding strategy problem of the DSA. Integrating demand flexibility with DERs, [5] presents a bidding strategy optimisation model for DA market involvement, modelling uncertain parameters of DERs and market prices. In [6], a stochastic decision-making model of a DSA's bidding strategy for the DA market is developed, coordinating renewable resources, DR, and Electric Vehicles (EVs). Taking into account the price elasticity of the retail loads, [7] proposes a bidding strategy for the DSA of DR resources in the DA wholesale market. The bidding strategy of a small-scale microgrids aggregator for participation in the RT balancing market is introduced in [8]; their proposed model aims to achieve maximum returns while mitigating the effects of uncertainties. However, such studies, [5], [6], [7], [8], remain narrow in focus dealing only with one market.

Some research studies investigated the involvement in multiple energy markets in an attempt to improve the DSA's profitability. In a study investigating the bidding strategy of a DR aggregator participating in DA and balancing markets, Vahid-Ghavidel et al. [9] report that the information-gap decision theory could be employed to gain a guaranteed predefined profit, taking into

Manuscript received 5 December 2022; revised 16 March 2023; accepted 14 April 2023. Date of publication 19 April 2023; date of current version 13 December 2023. This work was supported by Science Foundation Ireland under Grant SFI/16/IA/4496. Paper no. TEMPR-00083-2022. *(Corresponding author: Mohammad Afkousi-Paqaleh.)*

Mohammad Afkousi-Paqaleh and Andrew Keane are with the School of Electrical and Electronic Engineering, University College Dublin, D04V1W8 Dublin, Ireland (e-mail: mohammad.afkousi-paqaleh@ucdconnect.ie; andrew.keane@ucd.ie).

Mohammad Jafarian is with the EirGrid Groupe, D04FW28 Dublin, Ireland (e-mail: mohammad.jafarian@ucd.ie).

Digital Object Identifier 10.1109/TEMPR.2023.3268440

© 2023 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

account consumers' and market prices' uncertainties. In [10], a model is provided to optimise an EV aggregator's DA and RT bidding strategy in the presence of price uncertainty. According to [11], the hybrid stochastic-interval approach could be employed to reflect the uncertainties associated with renewable resources and different market prices in the bidding strategy problem of hybrid power plants. In [12] a bidding strategy model of a DR aggregator participating in DA and RT markets is presented. The problem's uncertainties are represented by a set of scenarios based on historical data. Using satisficing theory principles, [13] models the customer behaviour for the portfolio design of a DR aggregator. The proposed model determines the optimal DR premium for participation in DA and RT markets. In all the studies reviewed so far, different markets are recognised as independent and the interdependency of markets is overlooked. In [14] a coordinated strategy for a DSA that aggregates wind power and compressed air energy storage to participate in DA, Intra-Day (ID), and balancing markets is presented. The correlation among prices of different markets is considered via a constant factor.

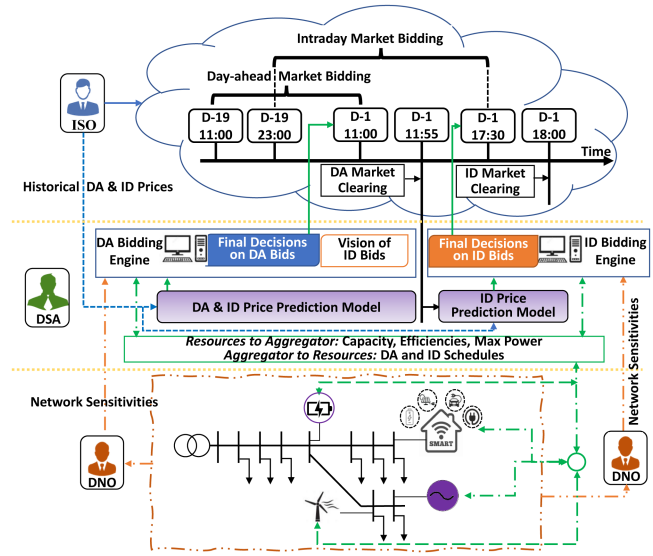


Fig. 1. Market structure and bidding strategy mechanism considering network constraints.

B. Contributions

To maximise their profits, DSAs are expected to engage in multiple electricity markets. In this study, a stochastic multi-step, two-stage bidding strategy problem is formulated and solved for DSAs participating in multiple markets. In this regard, to properly allocate their resources, DSAs should first identify and model the interdependencies of these markets as well as the price uncertainties. Although extensive research has been carried out on the bidding strategy problem of DSAs, no single study exists that models and integrates the interdependencies of multiple electricity markets into this problem. Merely considering different markets as independent is too simplistic. This assumption increases the risk of non-optimal allocation of resources in different markets, putting the DSA's profit at risk. Conversely, if the interdependence of markets is simplistically modelled by a constant correlation factor, there is the risk of under or over-estimating markets' dependence, thus diminishing the DSA's profit.

A framework based on Gaussian Process Regression (GPR) is proposed to model multiple electricity market prices. As GPR employs a kernel-based method, different data can be considered as explanatory parameters by specifying the respective kernel function. In addition, since GPR is a nonparametric Bayesian model, it is immune to data over-fitting, as is the case with linear models applied to non-linear data [15]. These characteristics allow GPR to capture and model 1) the interdependence between different electricity markets and 2) the uncertainties of market prices. Contrarily, neural networks can model the former but face unresolved challenges to model the latter [16], and ARIMA can only model the latter [17]. The GPR-based model complexity expands as the number of explanatory parameters increases. The Mutual Information Screening (MIS) technique is employed to limit the number of explanatory parameters to the proposed model by identifying the most relevant historical prices of all markets. In addition, a platform is developed to optimise the structure of the GPR by optimising and assigning different functions and parameters of the GPR. Finally, the trained model is implemented to generate price scenarios that incorporate the uncertainties and dependencies of various markets. These scenarios make it possible to optimise the bidding strategy problem of a price-taker DSA for participation in multiple markets.

In the bidding process, new data becomes available after each market clearing that could be exploited to increase the prediction accuracy for the upcoming markets. The DSA's bidding strategy problem is modelled as a multi-step two-stage problem to address this issue, wherein each step corresponds to a market. Following the clearing of each market, the resulting market prices are used to update the GPR-based model, thereby enhancing the accuracy of prediction for succeeding markets.

The results of solving the DSA bidding problem incorporating the multiple market price scenarios generated by the proposed GPR-based approach are compared to state-of-the-art price modelling approaches including, ARIMA, Lasso Estimated AutoRegressive (LEAR), and Long Short-Term Memory (LSTM) neural network.

II. Methodology

A. Market Structure

The DSA's involvement in the energy markets in this research follows the market rules provided in [18]. Nonetheless, the approaches outlined here could be applied to different markets with slight adjustments. In this section, only the DA and ID markets are modelled. However, the proposed approach could be easily expanded to encompass all existing markets. Fig. 1 schematically illustrates the structure of the bidding strategy problem and the mechanisms of the market. Three separate agents are illustrated: the Independent System Operator (ISO), the DSA, and the Distribution Network Operator (DNO). The

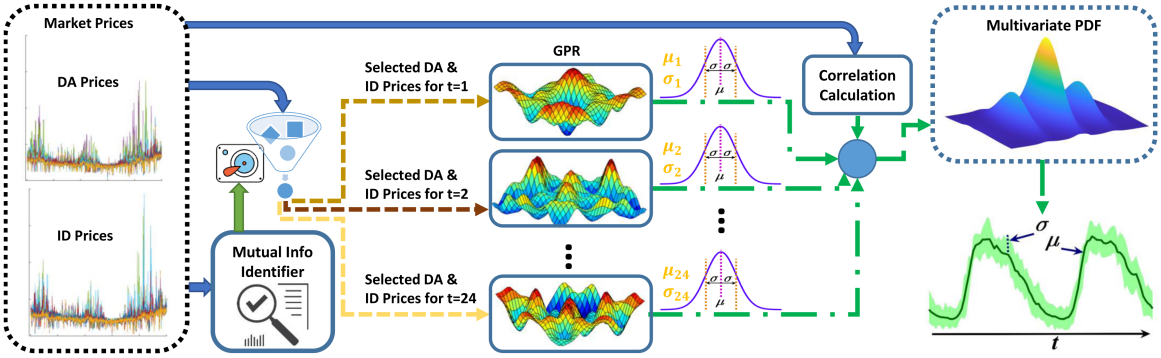


Fig. 2. Price prediction model.

ISO publishes historical market prices and releases market results following each market clearing. The DNO is responsible for ensuring that the electricity service is not imperilled while having limited monitoring infrastructure.

The DSA must submit its offers to tomorrow's DA market before 11:00. At this step, the DSA should have decided on its final DA bids and its vision of ID bids. After the DA market clearing, DSA leverages the DA market results to improve its prediction of ID prices. Finally, the DSA submits its bids for participation in the ID market. The DSA should strategically allocate its resources to different markets using a mathematical formulation that considers its resources' characteristics and price scenarios. Moreover, the DNO's provision of representative data enables DSA to build a set of equations with low computational cost to effectively construct constraints that reflect the bids' effect on the network's operational constraints.

B. Price Prediction Model

The DSA should strategise its offerings in the energy markets. To achieve this, DSA should first predict the DA and ID market prices using the ISO-supplied data. The price prediction could be accomplished using a set of scenarios that integrates the interdependencies between the two markets and represent the volatility of market prices. The proposed model for predicting prices is depicted in Fig. 2. A model based on GPR is developed to capture and integrate the interdependencies between markets to improve the accuracy of predictions. Owing to a kernel method, the GPR can model the interrelation dependency between different market prices as it accepts different types of explanatory parameters by configuring the kernel function of each parameter [19]. Being able to capture the input-output mapping from empirical data and to quantify the uncertainties of the outcome, GPR has been applied for short-term power forecast of photovoltaic [20], capacity estimation of lithium-ion batteries [21], and probabilistic load forecasting [22].

GPR is a nonparametric Bayesian model where its complexity grows as data size inflates. Therefore, GPR-based model training becomes computationally intensive when all historical prices across all markets are considered. The MIS technique is applied to identify the most relevant historical prices across all markets regarding the price to be predicted to make the proposed model computationally efficient. The proposed approach consists of two phases. First, the model is trained using historical data of different markets that contain the most amount of information, identified by MIS. Next, the trained model is implemented to predict future prices. In the implementation phase, the same structure of explanatory parameters (with new data) is given to the trained GPR-based model to predict the desired market prices.

A separate GPR-based model is developed for predicting the price of each hour of the day (and hence in total, 24 models). The dependent parameter obtained from the GPR-based model for each hour is a Gaussian distribution of the price at that hour. These Gaussian distributions are combined by employing Cholesky decomposition to generate a multivariate PDF of the market prices, considering the correlations between different hours. Finally, this multivariate PDF is used to generate the 24-hour price scenarios. The dependence of electricity prices across markets is described as scenario trees. Therefore, each DA market price scenario will have several ID price offspring. In other words, the ID price scenarios represent prospective ID market prices if that DA market price materialises. A high number of scenarios are generated and later reduced to thoroughly model market prices. In the first step (before DA market closure), the GPR-based model is used to create price scenarios for both markets, while in the second step (after DA market clearing), the GPR-based model is updated with actual DA prices to improve the ID price prediction accuracy.

1) Mutual Information Screening Technique: Mutual Information (MI) is a measure of mutual dependency between two random parameters. Specifically, it reveals how much information can be gleaned for one random parameter from observing another. Using the MIS technique, historical prices of all markets containing the most information about the to-be-predicted market price are identified. The primary distinction between correlation and the MIS technique is that correlation measures linear dependence, whereas the MIS technique measures generic dependence (including non-linear reliance). Suppose (X, Y) are two uncertain parameters (two market prices in this study) over the space $\mathcal{X} \times \mathcal{Y}$, with the joint probability distribution function of $\Pi_{(X,Y)}$, and the marginal distributions of Π_X and Π_Y. The

MI of the pair is:

$$MI(X;Y) = D_{KL}(\Pi_{(X,Y)} \| \Pi_X \otimes \Pi_Y) \tag{1}$$

where $D_{KL}(A\|B)$ is the Kullback–Leibler divergence, known as relative entropy, that measures the statistical distance of probability distribution A and reference probability distribution B. $\otimes$ indicates the tensor product. For discrete parameters, MI is defined as:

$$MI(X;Y) = \sum_{y\in Y}\sum_{x\in X}\left[\Pi_{(X,Y)} \log\left(\frac{\Pi_{(X,Y)}}{\Pi_X \Pi_Y}\right)\right] \tag{2}$$

MI could also be defined in terms of entropy. Denoting marginal entropies of X and Y by $H(X)$ and $H(Y)$, and joint entropy by $H(X,Y)$, MI could equivalently be expressed as:

$$MI(X;Y) \triangleq H(X) + H(Y) - H(X,Y) \tag{3}$$

The entropy of an uncertain parameter, $H(X)$, is the average degree of unpredictability and uncertainty inherent in the parameter's potential outcomes and is expressed as follows:

$$H(X) = -\sum_{x\in\mathfrak{X}} \Pi(X=x)\log\left(\Pi(X=x)\right) \tag{4}$$

The joint entropy, $H(X,Y)$, quantifies the uncertainty when two uncertain parameters are evaluated jointly.

$$H(X,Y) = -\sum_{x\in\mathfrak{X}}\sum_{y\in\mathfrak{Y}}[\Pi(X=x,Y=y) \log\left(\Pi(X=x,Y=y)\right)] \tag{5}$$

In other words, $H(X)$ is the degree of uncertainty of X, while $MI(X;Y)$ is the degree to which knowing Y reduces X's uncertainty. A larger MI value signifies a greater reduction in uncertainty of X, which corresponds to Y being more dependable. The unit of MI in this study is Shannon (bit).

2) Gaussian Process Regression: The GPR model describes the relationship between multivariate explanatory parameters (historical prices of different markets) and a scalar-dependent parameter (market price to be predicted). GPR uses the Gaussian process and Bayesian inference to estimate the nonlinear and nonparametric model between explanatory parameters and a dependent parameter [15]. Let $\Im_{train} = \{(\mathfrak{X}_n, \mathrm{y}_n)|n = 1,\ldots,N\}$ be the training data set, and $\Im_{test} = \{(\mathfrak{X}^*_m, \mathrm{y}^*_m)|m = 1,\ldots,M\}$ be the test data set, where $\mathfrak{X}_n$ is a vector, with the elements of $(\mathrm{x}^1_\mathrm{n}, \mathrm{x}^2_\mathrm{n}, \ldots, \mathrm{x}^\mathrm{D}_\mathrm{n})$ (N being the number of samples of explanatory parameters and D being the dimension of each sample of explanatory parameters) and y_n a scalar dependent parameter for the nth sample of explanatory parameters. The explanatory parameters-dependent parameter relation model could be expressed as follows:

$$\mathrm{y_n} = f(\mathfrak{X}_n) + \epsilon_n \tag{6}$$

where $f(\mathfrak{X})$ is a noise-free latent function and ϵ represents the noise with a Gaussian distribution $\mathcal{N}(0,\sigma_n^2)$.

Assuming $\mathbb{X} = [\mathfrak{X}_1;\mathfrak{X}_2;\ldots;\mathfrak{X}_N]_{N\times D}$ be the matrix of all samples of explanatory parameters, $\mathbb{Y}$ be a column vector of all dependent parameters, with the elements of $(\mathrm{y_1}, \mathrm{y_2}, \ldots, \mathrm{y_N})$, and $f(\mathbb{X})$ be a column vector of all latent function values of the training data, with the elements of $(f(\mathfrak{X}_1), f(\mathfrak{X}_2), \ldots, f(\mathfrak{X}_N))$; for all the training data set it can be stated that:

$$\mathbb{Y} = \boldsymbol{f}(\mathbb{X}) + \varepsilon \tag{7}$$

where ε denotes the noise vector with a Gaussian distribution $\mathcal{N}(0,\sigma_n^2\mathbf{I})$. The latent function $\boldsymbol{f}(\mathbb{X})$, is characterised by a multivariate Gaussian distribution:

$$\boldsymbol{f}(\mathbb{X}) \sim \mathcal{N}(0, K(\mathbb{X},\mathbb{X})) \tag{8}$$

where $K(\mathbb{X},\mathbb{X}) = K_N$ is the kernel matrix:

$$K(\mathbb{X},\mathbb{X}) = \begin{bmatrix} k(\mathfrak{X}_1,\mathfrak{X}_1) & \ldots & k(\mathfrak{X}_1,\mathfrak{X}_N) \\ k(\mathfrak{X}_2,\mathfrak{X}_1) & \ldots & k(\mathfrak{X}_2,\mathfrak{X}_N) \\ \ldots & \ldots & \ldots \\ k(\mathfrak{X}_N,\mathfrak{X}_1) & \ldots & k(\mathfrak{X}_N,\mathfrak{X}_N) \end{bmatrix} \tag{9}$$

The marginal distribution of $\mathbb{Y}$, $p(\mathbb{Y}|\mathbb{X})$, given that dependent parameters are normalised to a zero mean, could be obtained:

$$p(\mathbb{Y}|\mathbb{X}) = \int p(\mathbb{Y}|f(\mathbb{X}))p(f(\mathbb{X})|\mathfrak{X})df(\mathfrak{X}) \sim \mathcal{N}(0, K(\mathbb{X},\mathbb{X}) + \sigma_n^2\mathbf{I}) \tag{10}$$

The primary objective of GPR is to infer the distributions of test data-dependent parameters, y^*_m, given a fresh vector of explanatory parameters, $\mathfrak{X}^*_m$. The joint distribution of the training and test data for the function values $f(\mathfrak{X}_n)$ and $f(\mathfrak{X}^*_m)$ is given by:

$$p(f(\mathfrak{X}_n), f(\mathfrak{X}^*_m)|\mathfrak{X}_n,\mathfrak{X}^*_m) = \mathcal{N}\left(\begin{bmatrix} f(\mathfrak{X}_n) \\ f(\mathfrak{X}^*_m)\end{bmatrix}; 0, K_{N+M}\right) \tag{11}$$

$$K_{N+M} = \begin{bmatrix} K_N & K_{NM} \\ K_{MN} & K_M \end{bmatrix} \tag{12}$$

where K_M is the kernel matrix of test samples, and $K_{NM} = K^{tr}_{MN}$ (tr denoting transpose) are the kernel matrix between training and test samples. The joint distribution of y_n and y^*_m could be expressed as:

$$p\left(\mathrm{y}_n, \mathrm{y}^*_m|\mathfrak{X}_n,\mathfrak{X}^*_m\right) = \mathcal{N}\left(\begin{bmatrix} f(\mathfrak{X}_n) \\ f(\mathfrak{X}^*_m)\end{bmatrix}; 0, K_{N+M} + \sigma_n^2\mathbf{I}\right) \tag{13}$$

The predicted distribution of dependent parameters in the test data set can be determined by:

$$p(\mathrm{y}^*_m|\mathrm{y}_n,\mathfrak{X}_n,\mathfrak{X}^*_m) = \mathcal{N}(\mathrm{y}^*_m; \mu_m, K_{N+M} + \Sigma_m) \tag{14}$$

$$\mu_m = K_{MN}\left[K_N + \sigma_n^2\mathbf{I}\right]^{-1}\mathrm{y}^*_m \tag{15}$$

$$\Sigma_m = K_M - K_{MN}\left[K_N + \sigma_n^2\mathbf{I}\right]^{-1}K_{NM} \tag{16}$$

Almost $\mathcal{O}(N^3)$ computational cost is required to calculate the $[K_N + \sigma_n^2\mathbf{I}]^{-1}$ [19]. For practical application, the parameter vector $\vartheta = [K_N + \sigma_n^2\mathbf{I}]^{-1}\mathrm{y}_n$ is calculated in the training phase. The number of parameters in ϑ corresponds to the number of samples in the training data (N). Consequently, (15) could be

rewritten as follows (incurring only $\mathcal{O}(NM)$ computing cost):

$$\mu_m = K_{MN}\vartheta \tag{17}$$

To optimise the structure of the GPR-based model, a platform is developed which optimises the basis function, kernel function, kernel function's hyperparameters, kernel scale (when applicable), and sigma of GPR for each hour to improve performance. Depending on the characteristics of each hour, these parameters/functions are optimised or assigned. A function of each explanatory parameter could replace that parameter to better model the relationship between explanatory parameters and the dependent parameter. This function, known as the basis function, is affine, for which the coefficient and intercept are optimised. The kernel function is chosen from quadratic, exponential, squared exponential, Matern 5/2, Matern 3/2, or rational. The assigned kernel function's hyperparameters are optimised to achieve the best performance. Moreover, the platform searches among values within the range of explanatory parameters and values within the range of 0.0001 to 10 times the standard deviation of the dependent parameter to determine the kernel scale and sigma, respectively. Additionally, the explanatory parameters are normalised to a mean of zero and a standard deviation of one to eliminate the dependence on arbitrary scales in the explanatory parameters and enhance performance. The exact form of matrix representation used for the training phase will be used for the implementation (test) phase. The process mentioned above is done for each hour of the 24-hour ID market separately.

C. Problem Formulation

The bidding strategy for a DSA to partake in the DA and ID energy markets is modelled as a multi-step, two-stage stochastic optimisation problem aiming at profit maximisation. Each step represents a market. Prior to the closure of each market, the offerings in that market are considered first-stage variables, while the offerings in all future markets are considered second-stage variables. This problem is formulated as a Mixed Integer Linear Programming (MILP). Four types of resources for DSA are considered, including dispatchable DERs, wind farms, Battery Energy Storage (BES), and DRs.

1) First Step Objective Function, Before DA Market Closure: Initially, the DSA should determine bids for participation in the DA market and visions of bids in the ID market. The objective function of the DSA is profit (Υ) maximisation, which is comprised of the total profits from electricity markets minus the operational expenses, could be written as follows:

$$\text{Max } \Upsilon = \sum_{t=1}^{\mathcal{T}}\sum_{s=1}^{\mathcal{S}}\pi_s\left\{\left(P_t^{DA}\lambda_{s,t}^{DA} + P_{s,t}^{ID}\lambda_{s,t}^{ID} - \sum_{g=1}^{\mathcal{G}}P_{s,t}^{g}\beta^g\right)\tau - \sum_{b=1}^{\mathcal{B}}E_{s,t}^{b,tp}\beta^b - Cost_{s,t}^{DR}\right\} \tag{18}$$

where $\mathcal{T}$ is the number of time periods (hours) indexed by t, $\mathcal{S}$ is the number of scenarios indexed by s, and π_s is the probability of scenario s. Moreover, P_t^{DA} (MW) and $P_{s,t}^{ID}$ (MW) are the DSA's DA and ID offers, in scenario s and at time step t, respectively. $\lambda_{s,t}^{DA}$(€ /MWh) and $\lambda_{s,t}^{ID}$(€ /MWh) indicate the DA and ID prices of scenario s at time step t. On the second line of (18), $\mathcal{G}$ is the number of DERs indexed by g, and $\mathcal{B}$ is the number of BESs indexed by b. The power generation of DER g is denoted with $P_{s,t}^{g}$(MW), and $E_{s,t}^{b,tp}$(MWh) is the throughput energy of BES b, in scenario s and at time step t. The generation cost of DER g is signified by β^g(€ /MWh), while β^b(€ /MWh-throughput) denotes the degradation cost of BES b, and τ(h) symbolises the time step. Finally, the deployment cost of DR is represented by $Cost_{s,t}^{DR}$. It should be noted that the total offering at each hour of the DA market (P_t^{DA}) is a first-stage variable and constant in all scenarios, while the total offering at each hour of the ID market ($P_{s,t}^{ID}$) could vary over scenarios. P_t^{DA} and $P_{s,t}^{ID}$ are calculated using (19) and (20), respectively:

$$P_t^{DA} = \sum_{w=1}^{\mathcal{W}}P_{s,t}^{w,DA} + \sum_{d=1}^{\mathcal{D}}P_{s,t}^{d,DA} + \sum_{g=1}^{\mathcal{G}}P_{s,t}^{g,DA} + \sum_{b=1}^{\mathcal{B}}P_{s,t}^{b,DA} - \sum_{b=1}^{\mathcal{B}}L_{s,t}^{b,DA} \tag{19}$$

$$P_{s,t}^{ID} = \sum_{w=1}^{\mathcal{W}}P_{s,t}^{W,ID} + \sum_{d=1}^{\mathcal{D}}P_{s,t}^{d,ID} + \sum_{g=1}^{\mathcal{G}}P_{s,t}^{g,ID} + \sum_{b=1}^{\mathcal{B}}P_{s,t}^{b,ID} - \sum_{b=1}^{\mathcal{B}}L_{s,t}^{b,ID} \tag{20}$$

in the above equations, superscripts DA and ID represent the DA and ID markets. $\mathcal{W}$ is the number of wind farms indexed by w, and $\mathcal{D}$ is the number of DR participants indexed by d. Moreover, $P_{s,t}^{w}$, $P_{s,t}^{d}$, and $P_{s,t}^{b}$ are the power provision of the wind farm, DR resource, and BES in scenario s at hour t, respectively. The power used for charging BES at scenario s and period t is denoted by $L_{s,t}^{b}$. In Appendix VI-A, resource models and limitations are given.

Network Constraints Modelling: Untrammelled actions of DSAs could result in violation of network constraints and poor quality of service for customers. Therefore, DSAs should ensure that their bids do not violate network constraints.

$$\Omega_i^{Min} \leq \Omega_{i,t} \leq \Omega_i^{Max} \forall i, t \tag{21}$$

where Ω is the vector of network variables, i.e., nodal voltages and line currents, indexed by i, and t is the hour index. The maximum and minimum permissible ranges of network variables are denoted by Ω_i^{Max} and Ω_i^{Min}, respectively. The proposed bidding strategy takes the network operational constraints into account using the approach provided in [4], where the DNO does not interfere with the DSA's decision-making and there is no need for constant negotiation and data sharing between the two entities. This approach matches the state-of-the-art, and futuristic operation practices [23].

As DSAs have no access to network data, network constraints cannot be handled using a traditional OPF formulation [4]. Network sensitivity coefficients, which yield satisfactory results under different topologies and resource configurations [24], [25], [26], can be employed to measure the fluctuation in network variables ($\Omega_{i,t}$), due to changes in power generation/consumption by each resource. The approach to sensitivity analysis described in [26] is utilised because it has been demonstrated to be suitable for unbalanced feeders.

The DNO executes a series of calculations summarised in [4] and produces a linear regression that describes the variables of interest with a linear model. In other words, the degree to which the network variable $\Omega_{i,t}$ will change as a result of active and reactive power variation of resources is determined. Leveraging network variable estimates and sensitivity coefficients acquired from the DNO, DSAs can represent the magnitude of the network variables, $\Omega_{i,t}$, as a function of active and reactive power variation of resources. Consequently, these representations allow DSA to construct the set of network constraints, where each network variable $\Omega_{i,t}$ is bound within its permissible range. In contrast to conventional power flow equations, DSA does not require access to the network admittance matrix and thus is not responsible for network modelling (the DNO provides the network variable estimates and sensitivities).

2) Second Step Objective Function, After DA Market Clearing: Once the DA market has been cleared, the DSA leverages the DA market prices to update the GPR-based model. The updated GPR-based model is then used to create ID price scenarios. Moreover, in this step, the DSA's objective function will change. The DA offers should be fulfilled, however, considering the simultaneous delivery of DA and ID market offers during the RT operation of the grid, only the DA offers as a whole value (P_t^{DA}) are meaningful at this stage, and the exact utilisation of each resource in (19) and (20) is irrelevant and could change. Now, using all resources, the DSA should decide how to participate in the ID market, provided that the accepted DA offers are delivered.

$$\text{Max } \Upsilon^{ID} = \sum_{t=1}^{T}\sum_{s=1}^{S} \widehat{\pi}_s \left\{ \left(P_t^{ID}\widehat{\lambda_{s,t}^{ID}} - \sum_{g=1}^{\mathcal{G}} P_{s,t}^{g}\beta^{g} \right)\tau - \sum_{b=1}^{\mathcal{B}} E_{s,t}^{b,tp}\beta^{b} - Cost_{s,t}^{DR} \right\} \tag{22}$$

where $\widehat{\lambda_{s,t}^{ID}}$ and $\widehat{\pi}_s$ indicate the ID market prices and probability of scenario s obtained from the updated GPR-based model with actual DA prices. P_t^{ID} is determined using (23).

$$P_t^{ID} = \sum_{w=1}^{\mathcal{W}} P_{s,t}^{w} + \sum_{b=1}^{\mathcal{B}} P_{s,t}^{b} + \sum_{d=1}^{\mathcal{D}} P_{s,t}^{d} + \sum_{g=1}^{\mathcal{G}} P_{s,t}^{g} - \sum_{b=1}^{\mathcal{B}} L_{s,t}^{b} - P_t^{DA} \tag{23}$$

The ID offers are the accumulated utilisation of all resources minus the accepted DA offer. In other words, DSA may utilise the full capacity of its resources to participate in the ID market, provided that DA offers are delivered. The constraints provided in (25), (26), (30), (32), (34), (36), (38)-(41), and (21) curb the DSA's offer in the ID market.

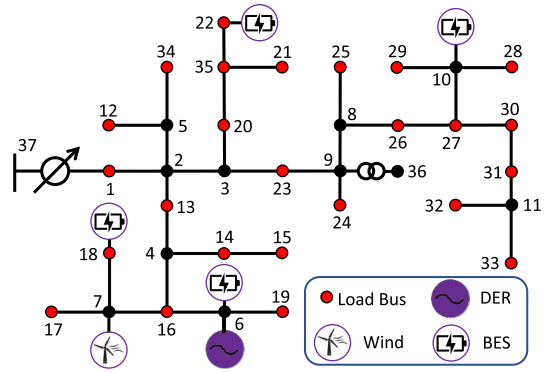


Fig. 3. The IEEE 37 node test feeder.

III. Case Data

A. Network Models

The IEEE 37 node [27] in Fig. 3, and IEEE 123-bus test feeder [28] are used as test systems. These systems operate at nominal voltages of 4.8 and 4.16 kV and include unbalanced loads. The daily load variations are modelled based on a representative of the Irish distribution system [29]. The IEEE 37-node test system is upgraded with a dispatchable DER, four storage units, and a wind farm. In addition, it is presumed that all load buses (red-highlighted in Fig. 3) engage in DR programs (see subsection III-C). The energy production cost of the 500 KW DER, β^g, is set at 60€ /MWh [30]. Four BESs with capacities 150, 200, 100, and 80 kWh are located on buses 6, 10, 18, and 22, respectively [31]. δ_{Conv}^{b} of all BESs is set at 25 per cent. The degradation cost of BES is modelled as 9€ /MWh-throughput, as determined by averaging the values presented in [32]. Moreover, the initial and final SOCs are fixed at 50%, while the minimum and maximum SOCs are set at 10% and 90%, respectively. Characteristics of the wind farm, located at bus 7, are given in subsection III-B3. The statutory voltage limits are 0.95 and 1.05 p.u. [33]. For comparative and reproducibility purposes, the IEEE 37 node test feeders' data and the optimization scripts in different cases for all approaches have been made available at the online repository [34].

B. Uncertainty Handling

1) Scenario Generation: For each day of the test data, 12,500 scenarios are generated. Each scenario has 72 parameters; the first 24 are the DA prices, the second 24 are the ID prices, and the final 24 represent the hourly wind generation. These scenarios are modelled as scenario trees, which means that for each DA price scenario, there are multiple ID price scenarios, and for each combination of DA and ID prices, there are several wind scenarios. After the DA market clearing, to determine the ID offers, the scenarios will have 48 parameters reflecting the ID prices and wind generation.

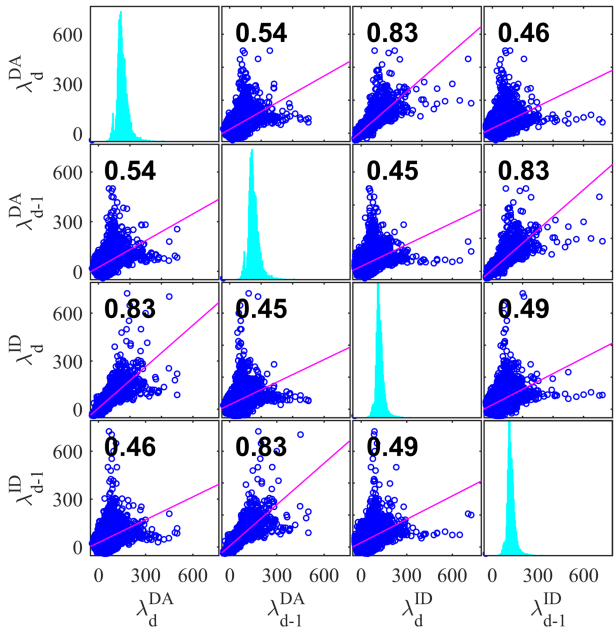


Fig. 4. Correlation matrix of DA and ID prices. Diagonal subplots depict the histogram of the distribution of market prices, and off-diagonal subplots depict a scatter plot of a pair of market prices, while the numbers show Pearson's correlation coefficients.

2) Price Scenarios: The historical DA and ID market prices are obtained from the Irish SEM Operator [35]. The price data from 1st October 2018 through 29th June 2021 are used. Of 1003 days, the first 802 days are used for training the model, and the remaining 201 days are considered test data. Each test day is analysed separately. It is worth noting that the Irish electricity market is highly volatile, and the price is negative in the early hours of some days due to excess wind generation. This volatility makes price prediction more challenging and complex, accentuating the necessity for price uncertainty modelling. The correlations of each day's DA (λ_d^{DA}) and ID (λ_d^{ID}) electricity market prices respecting each other and also the previous day's DA (λ_{d-1}^{DA}) and ID (λ_{d-1}^{ID}) market prices are illustrated in Fig. 4. It is interesting to note that the highest value of Pearson's linear correlation coefficients belongs to the correlation between λ_d^{DA} and λ_d^{ID}. This higher value of correlation shows that more information could be gleaned from the DA market clearing results regarding the ID market prices than the ID market prices of the previous day.

To generate price samples, first, the MIS technique is used to evaluate and identify the historical data that contains the highest amount of information regarding the next day's market prices. Fig. 5 depicts the amount of information included in the previous DA and ID prices regarding the ID prices. In Fig. 5(a), zero days ago means the DA prices of the same day after DA market clearing, and they have shown the highest degree of dependency. Similar to the findings of the correlation analysis, provided in Fig. 4, historical DA market prices contain more information regarding the expected ID prices than historical ID prices, confirming the need to model the interdependence of markets. As a threshold to restrict the number of explanatory parameters to the GPR-based, 0.3 b is used. Therefore, the

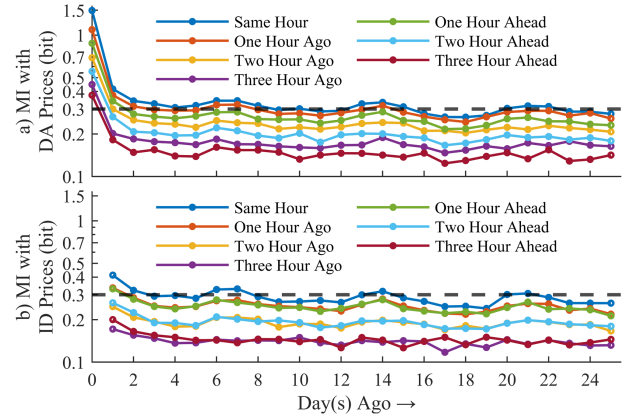


Fig. 5. MI of ID prices with historical DA and ID prices.

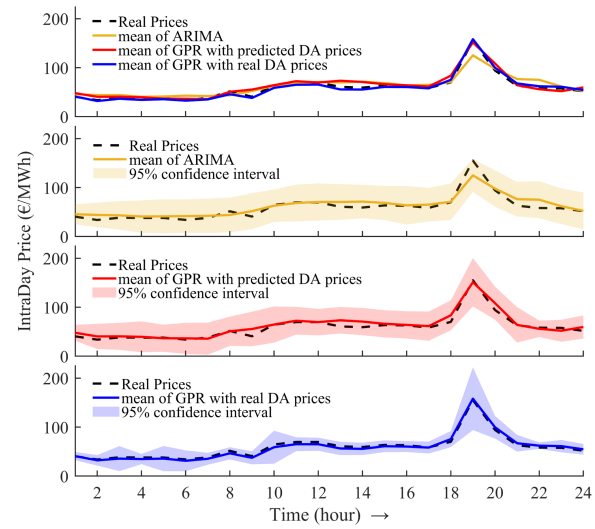


Fig. 6. ID price prediction results.

explanatory parameters to the GPR-based model consist of DA prices of all hours of the same day and all the preceding DA and ID prices that meet the threshold.

The GPR-based model is then trained using these prices (see II-B2). Next, employing the trained GPR-based model, hourly DA and ID price samples are generated, neglecting the correlations between hours. Cholesky decomposition is then used to incorporate correlations between various hours [33].

Fig. 6 displays the mean of predicted ID prices and the 95% confidence interval for ARIMA and GPR-based model for the first day of the test data. In this figure, "GPR with predicted DA prices" uses DA price predictions, whereas "GPR with actual DA prices" uses the real DA market prices following market clearing. The results of both GPR-based models show superior performance relative to the ARIMA. Moreover, when actual DA prices are used, the prediction accuracy dramatically increases, and the confidence interval shrinks.

3) Wind Generation Scenarios: The Beal Hill wind farm with a nominal power of 1.75 MW is modelled. The bias-corrected reanalysis approach [36] is utilised to determine the daily power generation of the wind farm across the study horizon. The acquired wind power generation is regarded as the mean

of a normal distribution, and a standard deviation of 2 per cent is used to reflect the volatility of the wind generation.

4) Scenario Reduction: Gams/Scenred2 [37] is used to reduce the sample size to 250, using a backward reduction approach. In the reduction phase, the probability of eliminated sample is added to the probability of the closest sample.

C. Demand Response

At each node and hour, it is estimated that 10% of the demand is willing to participate in the DR program. It is anticipated that customers will reduce their electricity usage in response to the incentive offered by DSA and that their demand will not shift to other hours. The self-elasticity of demand is set at −0.1 for all hours and buses, and consumers' electricity prices before and after partaking in the DR program are fixed at € 100/MWh [29]. The cost function of the DR program is approximated by ten segments. The DR is assumed to represent a variety of resources, including domestic load control, small storage, EVs, and electric heat pumps.

IV. Results and Discussion

In this section, the results of applying the proposed methodology are presented. Unless otherwise specified, all results pertain to the situation where network operational constraints are considered. In this regard, the results for a single day are first reported and discussed. The effect of considering network operational limitations is then evaluated. Next, the long-term analysis is provided, detailing the 201-day test data results. Finally, the implementation results of the proposed methodology on the IEEE 123-bus test system are given. All case studies are modelled and solved on a desktop computer with a 3-GHz i7 CPU and 16 GB of RAM using MATLAB R2020b and the MILP solver of Gurobi [38] in GAMS [37].

A. IEEE 37 Node Single-Day Results

As an example, the results of applying the proposed bidding strategy to the first day of test data are depicted in Fig. 7. This example illustrates how the proposed bidding strategy scheme allocates resources and determines optimal offerings in each market. These results reflect the DSA's bids following DA market clearing; thus, they are final. Fig. 7(a) shows the stochastic market prices derived using the GPR-based model. The final market offerings are depicted in Fig. 7(b). As noted before, the total resource utilisation for market participation should be a constant value across all scenarios, as shown in Fig. 7(b). This constraint is enforced by (19) for the DA market and (23) for the ID market. The DA bids result from solving (18), whereas the ID offers result from solving (22). Comparing the offers with hourly price data reveals that when DA prices are higher than ID prices, the resources are offered to the DA market and vice versa. Negative DA offers at hours 15 and 24 imply that the DSA purchases energy from the DA market to charge the BESs. The DSA also procures energy for BESs charging during other hours; however, since this energy is subtracted from other bids, it does not result in an overall negative value.

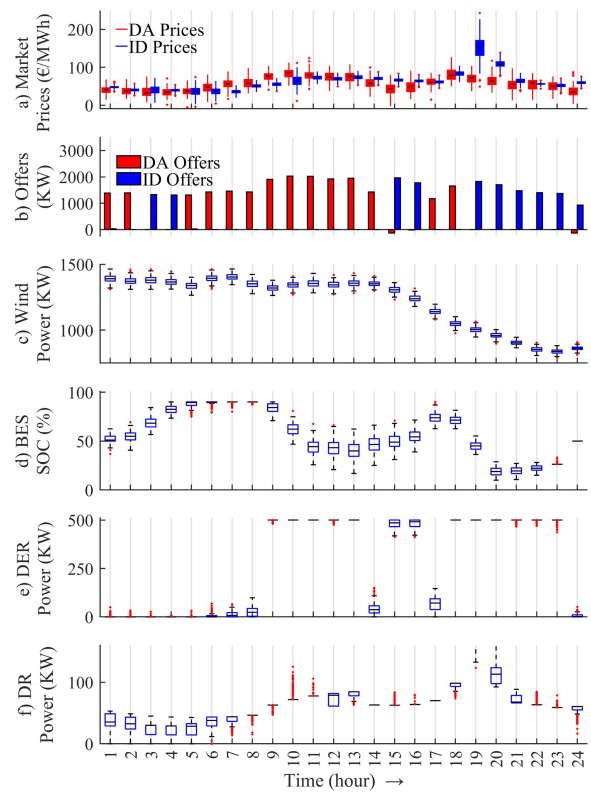
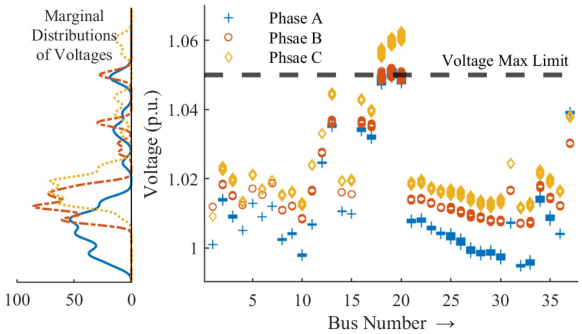


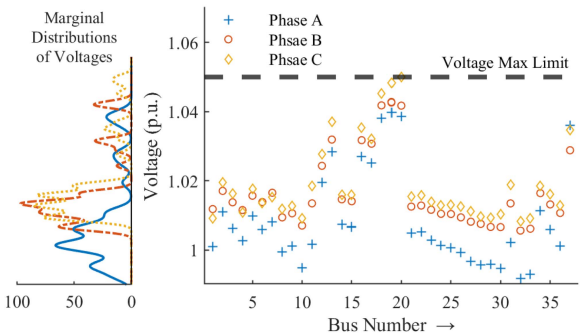
Fig. 7. The daily usage of resources.

Fig. 7 c illustrates wind power generation at various hours. This figure shows that the proposed approach incorporates the wind generation in the bids based on the model indicated in subsection III-B3. It depicts the adoption of a normal distribution with a standard deviation of a fixed per cent of the hourly mean generation to simulate the uncertainties associated with wind power generation. It is worth noting that wind spillage is zero in most scenarios and hours; however, there are a few scenarios in which wind spillage occurs at certain hours. The average utilisation of BESs for market participation is shown in Fig. 7(d). In the early hours of the day, when energy prices are relatively low, BESs are charged with energy bought from both the DA and ID markets. As electricity prices rise in the ensuing hours, the stored energy is sold back to the market, as indicated by a fall in the average SOC of BESs. Then, at hours 14-17, the BESs are recharged to be able to provide energy to the market at hours 18-20. In the last four hours, the average SOC of BESs increases as they are charged to satisfy the limit presented in (41). This figure demonstrates that the solution algorithm, across all scenarios, satisfies all constraints for the BES, including $SOC_{s,t}^{Min}$=0.1, $SOC_{s,t}^{Max}$=0.9, and $SOC_{s,t=24}^{b}$=0.5.

Fig. 7(e) displays the deployment of DER at various hours and scenarios. Since the DER has a fixed generating cost, it is beneficial to engage in the market when market prices are greater than β^g, and it is financially sensible to remain inactive when market prices are lower than β^g. During hours 10-11, for instance, market prices in all scenarios are higher than β^g; hence, the DER produces at its highest rate in all scenarios. However, when the difference between market prices and β^g is marginal (e.g., hours 7-8 and 17), DER is deployed in scenarios that



(a) Unconstrained Network Voltages



(b) Constrained Network Voltages

Fig. 8. Three phase bus voltages with and without network limits.

exhibit DA or ID market prices above β^g. Fig. 7(f) exhibits the accumulated utilisation of DR. As its application cost quadratically increases, DR is exploited moderately when market prices are relatively low. Alternatively, DR is implemented at a high level when market prices are high and can compensate for the expenses of the DR application.

B. Effect of Network Operational Constraints

Considering the sensitivities and magnitude of network variables provided by the DNO, the DSA should allocate its resources in a way that does not compromise the network's safety. Fig. 8 exhibits the three-phase node voltages of the network with and without voltage constraints. This figure represents the results for one hour ($t = 9$) across all scenarios, buses, and network phases. The marginal distribution of voltages shows the number of scenarios for each voltage level. According to Fig. 8(a), the voltages exceed the safe limit when no voltage limitations are imposed. As revealed by the marginal distribution of voltages, approximately 5 per cent of all nodal voltages exceed this constraint. In contrast, Fig. 8(b) shows the constrained network voltages, demonstrating that when the voltage limitations are applied, the proposed bidding strategy manages the resources in such a way that the network voltages across all scenarios, buses and phases are under the maximum limit. The marginal distribution of voltages above the permissible range is thus zero.

C. IEEE 37 Node Long-Term Analysis

Five case studies ($\mathcal{CS}$s) are defined and compared to thoroughly investigate the significance of the uncertainty and dependency modelling of different markets. $\mathcal{CS}1$ represents the

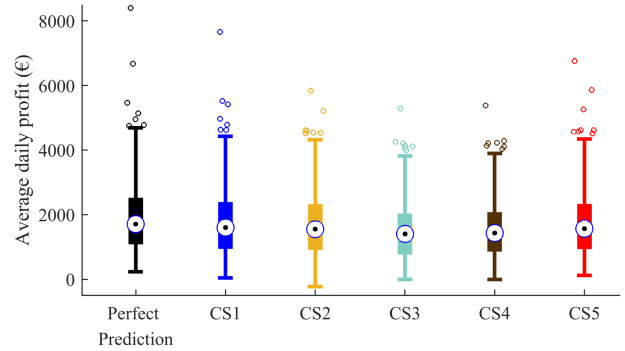


Fig. 9. Comparison of the profit of DSA in different cases over 201 days.

proposed methodology in which the DSA participate in both markets, taking into account the uncertainties and interdependencies. $\mathcal{CS}2$ is similar to $\mathcal{CS}1$ except that ARIMA is utilised instead of the GPR-based model for price prediction. Thus market uncertainties are modelled, but interdependence is neglected in this case. The distinction between $\mathcal{CS}3$ and $\mathcal{CS}1$ is that in $\mathcal{CS}3$, the DSA solely engages in the DA market. Similar to $\mathcal{CS}3$, the DSA in $\mathcal{CS}4$ only participates in one market; however, in this case, it is the ID market. $\mathcal{CS}3$ and $\mathcal{CS}4$ are defined to demonstrate how participation in only one market might vastly impact the DSA's revenue. The last case, $\mathcal{CS}5$, describes a situation in which a deterministic approach is used, and uncertainties associated with the bidding strategy problem are disregarded. In this case, the average of market prices and wind scenarios of $\mathcal{CS}1$ are given to the proposed methodology as a single scenario with the probability of 1.

Fig. 9 compares the profit of the DSA in these five cases and the perfect prediction case. The middle dot for each case represents the median value. The perfect prediction case represents the best possible solution with 100% accurate price predictions. The values depicted in this figure for the DSA's average daily profit are derived from solving (18). However, these values represent the actual profits of the DSA, which are calculated by substituting the DA and ID price scenarios with their actual values. On certain days in some cases, the profit of the DSA is negative, indicating that the price prediction was inaccurate and that the DSA incurred a loss. The average profit of the DSA in the perfect prediction case (the highest theoretically achievable profit) is € 1,957.5/day. This value is € 1,809.1 per day for the proposed methodology reported in $\mathcal{CS}1$. On average, the DSA could make € 1,737.9 per day if ARIMA is used to predict both DA and ID prices ($\mathcal{CS}2$). It is apparent from this figure that participating in only one market immensely reduces the profit as the average daily profit of the DSA is € 1,526.4 and € 1,580 per day in $\mathcal{CS}3$ and $\mathcal{CS}4$, respectively. If uncertainties are entirely disregarded, and fixed values are used for uncertain parameters ($\mathcal{CS}5$), the DSA can still earn an average of € 1,776/day.

The results demonstrate that modelling the uncertainties and interdependencies of markets significantly increases the DSA's profit. The observed 4.1% increase in the DSA's profit when the GPR-based model is used as opposed to ARIMA will equal € 14,311 over 201 days, which is considerable. The most interesting aspect of the results is that despite using a deterministic approach based on the mean of the price scenarios of $\mathcal{CS}1$ in $\mathcal{CS}5$

TABLE I
THE DEFINITION OF CASE STUDIES

Case study	DA Market	ID Market	Stochastic	Approach
$\mathcal{CS}1$	✓	✓	✓	GPR
$\mathcal{CS}2$	✓	✓	✓	ARIMA
$\mathcal{CS}3$	✓	✗	✓	GPR
$\mathcal{CS}4$	✗	✓	✓	GPR
$\mathcal{CS}5$	✓	✓	✗	GPR

TABLE II
RESULTS OF CASE STUDIES FOR IEEE 123-BUS TEST SYSTEM (€)

Case study	Avg. DA Profit	Avg. ID Profit	Avg. Daily Profit
Perfect Prediction	1105.46	1221.23	2326.69
$\mathcal{CS}1$	774.07	1215.80	1989.93
$\mathcal{CS}2$	769.24	1139.34	1908.61
$\mathcal{CS}3$	1732.73	-	1732.73
$\mathcal{CS}4$	-	1646.73	1646.73
$\mathcal{CS}5$	773.26	1151.48	1924.74

(interdependence modelled, uncertainty neglected), the DSA makes more profit than when a stochastic approach based on ARIMA is used in $\mathcal{CS}2$ (uncertainty modelled, interdependence neglected), confirming that modelling interdependence could be more important than modelling uncertainty.

D. IEEE 123-Bus Test Feeder

The proposed methodology is also implemented on the IEEE 123-bus test feeder. All data regarding this test feeder and the DERs' locations are obtained from [28]. However, the DERs are quadrupled in size. All other case data are the same as the previous system. Table II summarizes the results of the case studies outlined in Table I for this test system. The network operational variables are determined by employing the approach provided in [4] to ensure they are within the permissible range for this highly unbalanced feeder with overhead lines and underground cables.

In Table II, the DSA's profit in each market is given. The obtained results for this test feeder align with the results of the IEEE 37-node test system and corroborate the superiority of the proposed methodology. When markets' interdependence is modelled ($\mathcal{CS}1$), the average daily profit of the DSA rises by 4.3% as opposed to when it is neglected ($\mathcal{CS}2$). Comparing the average ID profit of the DSA in $\mathcal{CS}1$ and $\mathcal{CS}2$ with the perfect prediction case demonstrates that the proposed GPR-based methodology makes better decisions than ARIMA in the ID markets by leveraging the DA market clearing results. Moreover, the lack of diversity in DSA's resources in this test feeder (only dispatchable DERs) makes the profit of the DSA more dependent on the interdependence and uncertainty modelling of markets.

E. Comparison to State-of-The-Art Approaches

This subsection compares the bidding strategy results obtained by implementing the proposed GPR-based approach with state-of-the-art price prediction approaches, including ARIMA, LEAR and LSTM neural networks. The same structure of input data, as proposed in [39] is utilised for ARIMA. The LEAR approach is discussed in [40]. The LSTM parameters including the optimization algorithm, the number of epochs, gradient threshold, initial learning rate, and minimum batch size of LSTM are defined based on the values reported in [41]. The same structure of input data as those fed to GPR (the previous prices that have the highest amount of information identified by the MIS approach) is used in LSTM.

TABLE III
COMPARISON OF THE AVERAGE DAILY PROFIT (€) GAINED THROUGH DIFFERENT APPROACHES

Approach	37-Node Network	123-Node Network
Perfect Prediction	1957.474	2326.688
GPR	1809.056	1989.933
ARIMA	1737.944	1908.607
LSTM	1759.099	1894.708
LEAR	1793.311	1961.946

Table III presents the values for the DSA's profit for each test system with actual market prices. The results for both test systems demonstrate the efficacy of the proposed approach in better modelling the multiple market prices and thus improving the average daily profit of the DSA. Moreover, considering that the DA and ID price predictions are the same in these two test systems, the differences in the obtained solution to the bidding strategy problem are not only related to the efficacy of the price prediction approach but also related to the mix of resources available in each test system. Nonetheless, the proposed model shows superior performance in both cases.

V. CONCLUSION

DSAs should accurately model the market uncertainties and dependencies to engage in different energy markets. This work proposed a novel methodology based on the MIS technique and GPR to capture the uncertainties and dependence of different market prices. The DSA uses this data-driven uncertainty and interdependence handling model to determine resource allocation in multiple markets. Such resource allocation follows the relative price variation of markets, indicating the overall performance of the proposed methodology. The proposed bidding strategy was tested in various operating scenarios (201 days). The long-term analysis results indicate that the revenue will reduce if the DSA participates in only one market or fails to model the interdependencies between markets. The comparison of the bidding strategy results obtained utilising the proposed GPR-based approach with the state-of-the-art approaches unequivocally indicates the efficacy of the proposed markets-dependence modelling methodology.

It is envisaged that new markets of energy and services will be incorporated into the power system. It provides an opportunity for DSAs to increase their profits which could facilitate the integration of more DERs into the distribution network. This in turn defers or eliminates the need for distribution network upgrade investments, thereby enhancing social welfare. However, to exploit the multiple-market participation opportunity, the uncertainties and interdependence of these markets should be captured and factored into the bidding strategy problem as demonstrated by the obtained results. These points were

showcased and discussed at length for DA and ID markets in this paper. Moreover, as balancing markets are growing in importance and trading volumes across the world, the results of the application of the proposed method in these markets will be considered in future work.

Appendix

A. Resource Modelling

1) Wind Modelling: The total wind power generation is considered a stochastic parameter. Therefore, in each scenario, the upper bound is wind power available ($P^{w,Max}_{s,t}$) in that scenario as indicated by (24) and (25). Moreover, since the price could be negative at some hours, spillage ($P^{w,Spill}$) is introduced into the model to give the DSA the option of not providing its wind generation to the market. The spillage ($P^{w,Spill}_{s,t}$) should be less than the maximum available wind power at each scenario and each period as imposed by (26).

$$P^{w}_{s,t} = P^{w,DA}_{s,t} + P^{w,ID}_{s,t} \quad \forall w, s, t \tag{24}$$

$$P^{w}_{s,t} + P^{w,Spill}_{s,t} = P^{w,Max}_{s,t} \quad \forall w, s, t \tag{25}$$

$$P^{w,Spill}_{s,t} \leq P^{w,Max}_{s,t} \quad \forall w, s, t \tag{26}$$

2) DR Modelling: Incentive-based voluntary DR programs model as proposed in [29] is used to formulate the customers' behaviour. The total incentive paid to the DR resources at each hour and each scenario is calculated as follows:

$$Cost^{DR}_{s,t} = \sum_{d=1}^{\mathcal{D}} Inc^{d}_{t}\, P^{d}_{s,t}\tau \quad \forall d, s, t \tag{27}$$

$$P^{d}_{s,t} = L^{d}_{t} - L^{d,DR}_{s,t} \quad \forall d, s, t \tag{28}$$

where Inc^{d}_{t} (€ /MWh) is the incentive offered to the customers for each MWh reduction in their typical demand pattern. L^{d}_{t} and $L^{d,DR}_{t}$ are the demands (MW) of dth DR resource before and after DR programs, respectively. DR expenses are modelled using the concept of price elasticity of demand and linearised as proposed in [29]. The DR provision in DA and ID markets is bounded by the upper capacity of DR resources ($P^{d,Max}_{t}$).

$$P^{d}_{s,t} = P^{d,DA}_{s,t} + P^{d,ID}_{s,t} \quad \forall\, d, s, t \tag{29}$$

$$P^{d}_{s,t} \leq P^{d,Max}_{t} \quad \forall\, d, s, t \tag{30}$$

In the proposed DR program, the DSA is only concerned about the aggregate demand reduction of customers when they are asked to provide DR, thus, there is no need for detailed behind-the-meter data sharing. The DR participation in this paper is in line with the approach provided in [42] to preserve the privacy of the customers.

3) DER Modelling: DERs are deemed dispatchable generators. Equations (31) and (32) restrict the generation of DERs; while the former calculates the total offered generation of DERs to both markets, the latter caps the DERs' generation by their maximum limits ($P^{g,Max}$).

$$P^{g}_{s,t} = P^{g,DA}_{s,t} + P^{g,ID}_{s,t} \quad \forall g, s, t \tag{31}$$

$$P^{g}_{s,t} \leq P^{g,Max} \quad \forall g, s, t \tag{32}$$

4) BES Modelling: The constraints governing the operation and employment of BES are as follows:

$$P^{b}_{s,t} = P^{b,DA}_{s,t} + P^{b,ID}_{s,t} \quad \forall b, s, t \tag{33}$$

$$P^{b}_{s,t} \leq P^{b,Max} \quad \forall b, s, t \tag{34}$$

$$L^{b}_{s,t} = L^{b,DA}_{s,t} + L^{b,ID}_{s,t} \quad \forall b, s, t \tag{35}$$

$$L^{b}_{s,t} \leq P^{b,Max} \quad \forall b, s, t \tag{36}$$

$$E^{b,tp}_{s,t} = (P^{b}_{s,t} + L^{b}_{s,t})\tau \quad \forall b, s, t \tag{37}$$

$$P^{b,Max} = \delta^{b}_{Conv} E^{b,Max} \quad \forall b, s, t \tag{38}$$

$$SOC^{b}_{s,t} = SOC^{b}_{s,t-1} + \left(L^{b}_{s,t}\eta^{b}_{ch,Conv} - \frac{P^{b}_{s,t}}{\eta^{b}_{dc,Conv}} \right)\tau \quad \forall b, s, t \tag{39}$$

$$SOC^{b,Min}_{s,t} \leq SOC^{b}_{s,t} \leq SOC^{b,Max}_{s,t} \quad \forall b, s, t \tag{40}$$

$$SOC^{b}_{s,t=\mathcal{T}} = SOC^{b}_{s,t=0} \quad \forall b, s \tag{41}$$

The throughput energy of each BES at each scenario and time step is calculated using (37). The converter capacity, $P^{b,Max}$, bounds the total power provision or demand of each BES in all markets as indicated in (33)-(36). The converter capacity, which limits the maximum power transfer between the BES and the network at any moment, is modelled as a portion, δ^{b}_{Conv}, of the BES energy capacity, $E^{b,Max}$, as expressed in (38). The SOC of each BES at each time step of each scenario is calculated using (39) incorporating charging, $\eta^{b}_{ch,Conv}$, and discharging, $\eta^{b}_{dc,Conv}$ efficiencies. Moreover, in all times and scenarios, the SOC of BESs should be within their acceptable range (40). It should be noted that $L^{b}_{s,t}$ and $P^{b}_{s,t}$ are modelled on the network side of the connection. Considering the converter's efficiencies, charging and discharging at the same hour in the same market is not profitable. Therefore, the model does not need to enforce such a constraint. To this end, only charging and discharging at the same hour in different markets should be modelled. Therefore, using this concept, the number of binary variables, i.e., U, to prevent the charging and discharging at the same hour is halved using the following equations:

$$L^{b,DA}_{s,t} \leq P^{b,Max} U^{b,DA}_{s,t} \quad \forall b, s, t \tag{42}$$

$$P^{b,ID}_{s,t} \leq P^{b,Max}(1 - U^{b,DA}_{s,t}) \quad \forall b, s, t \tag{43}$$

$$L^{b,ID}_{s,t} \leq P^{b,Max} U^{b,ID}_{s,t} \quad \forall b, s, t \tag{44}$$

$$P^{b,DA}_{s,t} \leq P^{b,Max}(1 - U^{b,ID}_{s,t}) \quad \forall b, s, t \tag{45}$$

Acknowledgment

The opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Science Foundation Ireland.

REFERENCES

[1] J. Iria, F. Soares, and M. Matos, "Optimal supply and demand bidding strategy for an aggregator of small prosumers," *Appl. Energy*, vol. 213, pp. 658–669, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261917312758

[2] F. Wang, X. Ge, K. Li, and Z. Mi, "Day-ahead market optimal bidding strategy and quantitative compensation mechanism design for load aggregator engaging demand response," *IEEE Trans. Ind. Appl.*, vol. 55, no. 6, pp. 5564–5573, Nov./Dec. 2019.

[3] A. Hajebrahimi, I. Kamwa, M. M. A. Abdelaziz, and A. Moeini, "Scenario-wise distributionally robust optimization for collaborative intermittent resources and electric vehicle aggregator bidding strategy," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 3706–3718, Sep. 2020.

[4] V. Rigoni, D. Flynn, and A. Keane, "Coordinating demand response aggregation with LV network operational constraints," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 979–990, Mar. 2021.

[5] M. Di Somma, G. Graditi, and P. Siano, "Optimal bidding strategy for a DER aggregator in the day-ahead market in the presence of demand flexibility," *IEEE Trans. Ind. Electron.*, vol. 66, no. 2, pp. 1509–1519, Feb. 2019.

[6] P. Afzali, M. Rashidinejad, A. Abdollahi, and A. Bakhshai, "Risk-constrained bidding strategy for demand response, green energy resources, and plug-in electric vehicle in a flexible smart grid," *IEEE Syst. J.*, vol. 15, no. 1, pp. 338–345, Mar. 2021.

[7] M. Ostadijafari, R. R. Jha, and A. Dubey, "Demand-side participation via economic bidding of responsive loads and local energy resources," *IEEE Open Access J. Power Energy*, vol. 8, pp. 11–22, 2021.

[8] W. Pei, Y. Du, W. Deng, K. Sheng, H. Xiao, and H. Qu, "Optimal bidding strategy and intramarket mechanism of microgrid aggregator in real-time balancing market," *IEEE Trans. Ind. Informat.*, vol. 12, no. 2, pp. 587–596, Apr. 2016.

[9] M. Vahid-Ghavidel, N. Mahmoudi, and B. Mohammadi-Ivatloo, "Self-scheduling of demand response aggregators in short-term markets based on information gap decision theory," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 2115–2126, Mar. 2019.

[10] H. Yang, S. Zhang, J. Qiu, D. Qiu, M. Lai, and Z. Dong, "CVaR-constrained optimal bidding of electric vehicle aggregators in day-ahead and real-time markets," *IEEE Trans. Ind. Informat.*, vol. 13, no. 5, pp. 2555–2565, Oct. 2017.

[11] H. Khaloie, A. Anvari-Moghaddam, N. Hatziargyriou, and J. Contreras, "Risk-constrained self-scheduling of a hybrid powerplant considering interval-based intraday demand response exchange market prices," *J. Cleaner Prod.*, vol. 282, 2021, Art. no. 125344. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652620353890

[12] R. Henríquez, G. Wenzel, D. E. Olivares, and M. Negrete-Pincetic, "Participation of demand response aggregators in electricity markets: Optimal portfolio management," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4861–4871, Sep. 2018.

[13] F. Salah, R. Henríquez, G. Wenzel, D. E. Olivares, M. Negrete-Pincetic, and C. Weinhardt, "Portfolio design of a demand response aggregator with satisficing consumers," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2475–2484, May 2019.

[14] S. Ghavidel, M. J. Ghadi, A. Azizivahed, J. Aghaei, L. Li, and J. Zhang, "Risk-constrained bidding strategy for a joint operation of wind power and CAES aggregators," *IEEE Trans. Sustain. Energy*, vol. 11, no. 1, pp. 457–466, Jan. 2020.

[15] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning., Ser. Adaptive Computation and Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

[16] B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and scalable predictive uncertainty estimation using deep ensembles," in *Advances in Neural Information Processing Systems*, vol. 30. Red Hook, NY, USA: Curran Associates, Inc., 2017.

[17] R. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. Melbourne, VIC, Australia: OTexts, 2021.

[18] Single Electricity Market Committee, "The single electricity market (SEM), trading and settlement code V. 26.0," Single Electricity Market Operator, Market Rules, May 2022. [Online]. Available: https://www.sem-o.com/rules-and-modifications/balancing-market-modifications/market-rules/Part-A-TSC.docx

[19] T. Koriyama, T. Nose, and T. Kobayashi, "Statistical parametric speech synthesis based on gaussian process regression," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 2, pp. 173–183, Apr. 2014.

[20] H. Sheng, J. Xiao, Y. Cheng, Q. Ni, and S. Wang, "Short-term solar power forecasting based on weighted Gaussian process regression," *IEEE Trans. Ind. Electron.*, vol. 65, no. 1, pp. 300–308, Jan. 2018.

[21] R. R. Richardson, C. R. Birkl, M. A. Osborne, and D. A. Howey, "Gaussian process regression for insitu capacity estimation of lithium-ion batteries," *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 127–138, Jan. 2019.

[22] D. Cao et al., "Robust deep Gaussian process-based probabilistic electrical load forecasting against anomalous events," *IEEE Trans. Ind. Informat.*, vol. 18, no. 2, pp. 1142–1153, Feb. 2022.

[23] National Grid ESO, "Power potential (transmission & distribution interface 2.0) project close down report," U.K. Power Networks, London, U.K., Tech. Rep., 2021. [Online]. Available: https://www.nationalgrideso.com/document/205371/download

[24] G. Valverde and T. Van Cutsem, "Model predictive control of voltages in active distribution networks," *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 2152–2161, Dec. 2013.

[25] Z. Zhang, L. F. Ochoa, and G. Valverde, "A novel voltage sensitivity approach for the decentralized control of DG plants," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1566–1576, Mar. 2018.

[26] V. Rigoni and A. Keane, "Estimation of voltage sensitivities in low voltage feeders with photovoltaics," in *Proc. IEEE PES Innov. Smart Grid Technol. Conf. Europe*, 2018, pp. 1–6.

[27] W. H. Kersting, "Radial distribution test feeders," *IEEE Trans. Power Syst.*, vol. 6, no. 3, pp. 975–985, Aug. 1991.

[28] M. Jafarian, A. Soroudi, and A. Keane, "Distribution system topology identification for DER management systems using deep neural networks," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2020, pp. 1–5.

[29] M. Afkousi-Paqaleh, V. Rigoni, and A. Keane, "Network-constrained decentralised bidding strategy of DR aggregators considering customer behaviour," in *Proc. IEEE PES Innov. Smart Grid Technol. Europe*, 2021, pp. 1–6.

[30] A. Nouri, A. Soroudi, and A. Keane, "Strategic scheduling of discrete control devices in active distribution systems," *IEEE Trans. Power Del.*, vol. 35, no. 5, pp. 2285–2299, Oct. 2020.

[31] M. M. Esfahani, A. Hariri, and O. A. Mohammed, "A multiagent-based game-theoretic and optimization approach for market operation of multimicrogrid systems," *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 280–292, Jan. 2019.

[32] G. He, S. Kar, J. Mohammadi, P. Moutis, and J. F. Whitacre, "Power system dispatch with marginal degradation cost of battery storage," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3552–3562, Jul. 2021.

[33] A. Nouri, M. Jafarian, and A. Keane, "Voltage restoration after unforeseen disturbances in weakly observable distribution systems," *IEEE Trans. Power Syst.*, vol. 37, no. 4, pp. 3063–3076, Jul. 2022.

[34] M. Afkousi-Paqaleh, M. Jafarian, and A. Keane, "Case data, and optimization codes for "modelling the interdependence of multiple electricity markets in the distribution system aggregator bidding," 2023. [Online]. Available: https://figshare.com/s/a0406844bc8ff2f0b163

[35] Single Electricity Market Operator, "Market data," 2022. [Online]. Available: https://www.semopx.com/market-data/market-results/

[36] I. Staffell and S. Pfenninger, "Using bias-corrected reanalysis to simulate current and future wind power output," *Energy*, vol. 114, pp. 1224–1239, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360544216311811

[37] GAMS Development Corporation, GAMS–Documentation, 2021. [Online]. Available: https://www.gams.com/latest/docs/gams.pdf

[38] G. Optimization, LLC, "Gurobi optimizer reference manual," 2022. [Online]. Available: https://www.gurobi.com

[39] J. Contreras, R. Espinola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1014–1020, Aug. 2003.

[40] J. Lago, G. Marcjasz, B. De Schutter, and R. Weron, "Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark," *Appl. Energy*, vol. 293, 2021, Art. no. 116983. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261921004529

[41] Z. Shao et al., "Modeling and forecasting the electricity clearing price: A novel BELM based pattern classification framework and a comparative analytic study on multi-layer BELM and LSTM," *Energy Econ.*, vol. 86, 2020, Art. no. 104648. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140988319304451

[42] S. Talari, M. Shafie-Khah, N. Mahmoudi, P. Siano, W. Wei, and J. P. Catalão, "Optimal management of demand response aggregators considering customers' preferences within distribution networks," *IET Gener., Transmiss. Distrib.*, vol. 14, no. 8, pp. 5571–5579, Dec. 2020. [Online]. Available: https://digital-library.theiet.org/content/journals/10.1049/iet-gtd.2020.1047