

Multi-objective robust dynamic VAR planning in power transmission grids for improving short-term voltage stability under uncertainties

ISSN 1751-8687

Received on 30th September 2017

Revised 19th December 2017

Accepted on 4th January 2018

doi: 10.1049/iet-gtd.2017.1521

www.ietdl.orgTong Han¹, Yanbo Chen¹✉, Jin Ma²¹School of Electrical and Electronic Engineering, North China Electric Power University, 102206, Beijing, People's Republic of China²School of Electrical and Information Engineering, University of Sydney, NSW, 2006, Australia

✉ E-mail: yanbochen2008@sina.com

Abstract: Modern power transmission grids are facing more and more critical short-term voltage stability problems. This study proposes a novel robust dynamic VAR planning approach for improving the short-term voltage stability level under uncertainties including the peak load level, the maximum proportion of dynamic load, fault clearing time and deviation of the actual capacity of dynamic VAR compensators from rated capacity when contingencies occur. The robust dynamic VAR planning problem is formulated as a multi-objective optimisation model with objectives including the investment cost, the expectation and robustness of the short-term voltage stability level. The complexity of the planning model is firstly reduced by selecting severe contingencies and potential buses, leading to a simplified multi-objective optimisation model. Latin hypercube sampling is then used for the uncertainty quantification. The simplified multi-objective optimisation model is then solved by the combination of a multi-objective evolutionary algorithm called ϵ -NSGAII and extreme learning machine-based surrogate modelling with adaptive training data sampling. This combination significantly reduces the unaffordable computing burden. Simulations are carried on the IEEE 39-bus system, illustrating that the authors proposed methodology is reliable with high computational efficiency, and offering decision-makers planning solutions with high mean performance and strong robustness with respect to the short-term voltage stability level.

1 Introduction

With the increasing proportion of air-conditioning load, growing use of capacitor banks for reactive compensation and more intensive use of transmission, modern power transmission grids are facing more and more critical short-term voltage stability problems. The root causes of some major grid blackouts, such as the northeast blackout of 2003 and the 2012 India blackout, causing power outage among millions of customers and significant financial loss, are related to short-term voltage instability [1, 2]. Thus the level of short-term voltage stability is vital to the security of power grids. Also, compared with other forms of stability, short-term voltage stability, with a period in the order of several seconds, needs fast controls to enhance [3, 4].

Dynamic VAR compensators such as synchronous compensators and static synchronous compensators (STATCOMs) can efficiently improve the short-term voltage stability level by offering fast-acting dynamic reactive power support to power grids. The effectiveness, however, is heavily subject to the size, location, and type of dynamic VAR compensator [5, 6]. Thus the issue of dynamic VAR planning is of great significance when using dynamic VAR compensators to improve short-term voltage stability, and to which academia and industry have paid more and more attention.

There exist some studies addressing the issue of dynamic VAR planning for improving short-term voltage stability, modelling from various viewpoints and using different optimisation approaches. A method based on the concept of trajectory sensitivity is proposed to identify the location for dynamic VAR compensators to mitigate the short-term voltage instability problem [7]. Trajectory sensitivity analysis, although unable to determine the optimal size of dynamic VAR compensators, is widely used to select the location or potential location buses before optimising [8–12]. Also, then with optimisation theory, the dynamic VAR planning problem is formulated as a single- or multi-objective model.

In the framework of single-objective optimisation, the total cost of dynamic VAR compensators installed and the performance indices with respect to short-term voltage stability are modelled as the objective to be minimised and constraints, respectively [8, 9, 12, 13], or the inverse when the amount of dynamic VAR compensators to be installed is given [14]. Constraints can also include differential-algebraic equation (DAE) constraints for system dynamics, transient stability constraints, pre-disturbance steady state operation constraints, operational limit constraints and other pertinent constraints [8–10, 12–14]. As for the performance indices with respect to short-term voltage stability, voltage recovery time and duration of low voltage [8], or analogously the WECC criteria [9, 10], have been firstly used. Also, alternatively, an index based on voltage trajectory violation integral is then used to assess the short-term voltage stability level [13]. The empirical controllability covariance, indirectly estimating the improvement of the short-term voltage stability level after installing dynamic VAR compensators, is also used [14]. To solve the optimisation models, various optimisation approaches, e.g., branch-and-bound algorithm [8], control vector parameterisation approach [9], heuristic optimisation [13], Voronoi diagram approach integrating linear programming [10, 12] and mesh adaptive direct search algorithm [14], have been applied.

In the framework of multi-objective optimisation, the total cost of dynamic VAR compensators installed and short-term voltage stability level is both modelled as objectives to be optimised simultaneously [11, 15–17]. Compared with single-objective models of dynamic VAR planning offering only one optimal solution, multi-objective models can offer more choices to decision makers by getting a set of optimal solutions, called Pareto-optimal solutions. Some power grid performances related to dynamic VAR compensation such as the transient stability level can also be involved as an objective to make trade-offs [15, 16]. Also, multi-objective algorithms including the decomposition-based evolutionary algorithm and improved non-dominated sorting genetic algorithm have been applied to solve the model.

Due to the large scale of practical power transmission grids and especially the complexity of power grids' dynamics, it is generally indispensable to identify the most severe contingencies and the most potential location buses before optimising to reduce the complexity of the optimisation model. After obtaining values of the defined severity index of all credible contingencies, manmade selection approach [8, 11, 13], K-means clustering [15], and Pareto optimality based approach [16] have been applied to the selection of severe contingencies. Also, for the selection of local sensitivity analysis [8, 9, 11, 16], engineering-based analysis [13], and global sensitivity analysis [15], have been applied. An exception is the work [17], which directly executes optimisation without a selection of contingencies and location buses by employing high-performance computing clusters.

Although many efforts have been devoted to the issue of dynamic VAR planning for improving short-term voltage stability, all existing studies formulate the problem as deterministic models and there is still no work addressing uncertainties in the planning model. However, some factors in dynamic VAR planning, having a significant impact on planning results, are essentially stochastic or probabilistic. For example, the accurate forecasting of the peak load level in the aimed year which greatly effects short-term voltage stability of power grids is a formidable task due to a variety of uncertainties [18, 19]. Thus the optimal solutions obtained with a deterministic peak load level can be suboptimal, lack robustness to the uncertainty in the peak load level or even infeasible in the actual situation. To this problem, this paper aims at unprecedently addressing the uncertainties in dynamic VAR planning for improving short-term voltage stability under the framework of multi-objective optimisation. The main contributions of this study are threefold: (i) we analyse and model the uncertainties in dynamic VAR planning. (ii) For addressing dynamic VAR planning problem under uncertainties, we propose a multi-objective robust dynamic VAR planning model with objectives including the investment cost, the expectation, and robustness of the short-term voltage stability level. (iii) To ease the computation burden caused by power system dynamics and uncertainty variables, an approach combining surrogate modelling technologies with adaptive training data sampling and a multi-objective optimisation algorithm called ϵ -NSGAII, is proposed to efficiently solve the robust planning model. Note that in the following part of this paper, the term dynamic VAR planning means dynamic VAR planning for improving short-term voltage stability, for short.

The rest of this paper is organised as follows. The classical multi-objective dynamic VAR planning model is given in Section 2. Uncertainties in dynamic VAR planning is also analysed in this section. In Section 3, the multi-objective robust optimisation model of dynamic VAR planning is formulated. The solving approach of the proposed model is given in Section 4. Section 5 shows the test results of the proposed multi-objective robust dynamic VAR planning approach applied to the IEEE 39-bus system. Section 6 concludes this paper.

2 Short reviews of classical multi-objective dynamic VAR planning model and related uncertainties

2.1 Classical multi-objective dynamic VAR planning model

Take STATCOMs' planning as an example hereinafter, but the planning of other dynamic VAR compensators is the analogue. Note that the choosing of compensators' type is not taken into account in this study. The classical multi-objective dynamic VAR planning model in the prior work, making trade-offs between the total investment cost and short-term voltage stability level, is given in [11, 15, 16]

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{F}' = [F_1(\mathbf{x}), F_2(\mathbf{x})], \\ \text{s.t.} \quad & \begin{cases} \mathbf{h}'_1(\mathbf{x}) = \mathbf{0}, \\ \mathbf{h}'_2(\mathbf{x}) \leq \mathbf{0}, \\ S, \end{cases} \end{aligned} \quad (1)$$

where \mathbf{x} is the vector of decision variables and $\mathbf{x} = (x_i)^{n_c \times 1} \subset (\mathbb{Z}_0^+)^{n_c \times 1}$ with $i \in \Lambda$, x_i denoting the rated capacity of STATCOMs installed at bus i , Λ being the set of initial candidate buses to install STATCOMs, and n_c representing the number of buses in Λ . The sub-objectives F'_1 and F'_2 represent the total investment cost of STATCOMs to be installed and short-term voltage stability level, respectively. Also, $\mathbf{h}'_1(\mathbf{x}) = \mathbf{0}$ and $\mathbf{h}'_2(\mathbf{x}) \leq \mathbf{0}$ represent equality constraints such as DAE constraints, and inequality constraints, respectively. The term S denotes a certain forecasted operating condition generally with a peak load level and maximum proportion of dynamic load in the aimed year.

The sub-objective F'_2 in (1) can be written as

$$F'_2(\mathbf{x}) = \sum_{\Theta} p_k \psi'(\mathbf{x}, k) \quad (2)$$

with Θ denoting the set of credible contingencies, p_k denoting the occurrence probability or the weight of contingency k , and ψ' being an index of the short-term voltage stability level under a certain contingency. The lower values of ψ' represent the higher short-term voltage stability level. Also, before optimising, Θ and Λ should be, respectively, replaced by the set Θ_s containing severe contingencies and the set Λ_c containing potential location buses according to selection results to obtain the simplified optimisation model.

2.2 Uncertainties in dynamic VAR planning

Classical approaches all model the multi-objective dynamic VAR planning problem as a deterministic optimisation model shown in (1). However, in fact, since dynamic VAR planning is a long-term planning problem, the forecast of parameters in S can be a major source of uncertainties [20]. And among them, the peak load level and maximum proportion of dynamic load greatly affect short-term voltage stability of power grids and thus have a substantial impact on planning results. Additionally, fault clearing time of contingencies in Θ_s , a determinant of short-term voltage stability, is also with uncertainties [21]. In addition to the above uncertainties from the power grids' interior, due to the operation requirements of power grids, the actual capacity of dynamic VAR compensators is probably not equal to rated capacity when contingencies occur. Thus the deviation of the actual capacity of dynamic VAR compensators from rated capacity is also a major source of uncertainties.

Based on the above analysis, we model the peak load level α_1 , the maximum proportion of dynamic load α_2 , fault clearing time α_3 and deviation of the actual capacity of dynamic VAR compensators from rated capacity when contingencies occur δ_i as uncertain variables, and the robustness of short-term voltage stability against them is considered. The quantification of these uncertainties is given below.

The uncertainty of the peak load level in the aimed year has been modelled by a normal distribution or an interval [19, 22]. Combining these two quantifications, we model the uncertainty of α_1 by the truncated normal distribution, and the probability density function (PDF) of α_1 is given by

$$f_a(\alpha_1 | \mu_1, \sigma_1, a_l, b_l) = \begin{cases} \frac{(1/\sigma_1)\phi((\alpha_1 - \mu_1)/\sigma_1)}{\Phi((b_l - \mu_1)/\sigma_1) - \Phi((a_l - \mu_1)/\sigma_1)}, & a_l \leq \alpha_1 \leq b_l, \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

where μ_1 , σ_1 , a_l and b_l denote the mean, standard deviation, lower tail and upper tail of the distribution, respectively. Also, $\phi(\xi)$ and $\Phi(\xi)$ are the PDF and cumulative distribution function (CDF) of the standard normal distribution, and they are given as

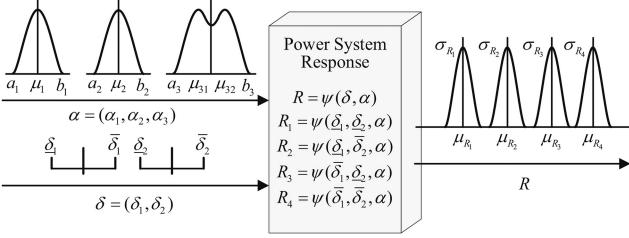


Fig. 1 Robustness estimation of the response R

$$\phi(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\xi^2\right), \quad (4)$$

$$\Phi(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\xi} \exp\left(-\frac{1}{2}\xi^2\right) d\xi. \quad (5)$$

Also, analogously, the uncertainty of α_2 is also modelled by a truncated normal distribution with $f_a(\alpha_2 | \mu_2, \sigma_2, a_2, b_2)$ as its PDF (see (6)) In theory, the uncertainty of fault clearing time α_3 can be modelled by a linear combination of two normal distributions [21]. However, in practice, fault clearing time can be neither infinitely short nor infinitely long. Thus in analogy to the above truncated normal distribution, by introducing lower tail a_3 and upper tail b_3 , the PDF of α_3 can be represented by (6), where μ_{31} and μ_{32} are the means of the two combined normal distributions, and δ_3 is the shared standard deviation of them.

Moreover, the uncertainty of δ_i is related to the control system of dynamic VAR compensators, coordinated voltage-control schemes [23], etc. There is, however, a lack of knowledge and information about the uncertainty of δ_i in the planning phase. Thus δ_i is an epistemic uncertainty variable and its uncertainty can be modelled as an interval $[\underline{\delta}_i, \bar{\delta}_i]$ with $\underline{\delta}_i = 0$ and $\bar{\delta}_i > 0$. Also, $\alpha_1, \alpha_2, \alpha_3$ and δ_i are assumed to be independent.

3 Optimisation model of multi-objective robust dynamic VAR planning

In this section, we formulate the dynamic VAR planning problem incorporating uncertainties in the framework of multi-objective optimisation analogous to (1), making trade-offs among the investment cost, and two statistical measures which are optimised simultaneously to ensure the mean performance and robustness of planning solutions under uncertainties. Objectives and constraints of the proposed optimisation model are described in detail below.

3.1 Objectives of the proposed optimisation model

Objectives of the proposed optimisation model for multi-objective robust dynamic VAR planning are given by

$$\min_{\mathbf{x}} \mathbf{F} = [F_1(\mathbf{x}), F_2(\mathbf{x}, \boldsymbol{\delta}, \boldsymbol{\alpha}), F_3(\mathbf{x}, \boldsymbol{\delta}, \boldsymbol{\alpha})], \quad (7)$$

where \mathbf{x} with the same meaning as that in (1) is the vector of decision variables. Also, $\boldsymbol{\delta}$ and $\boldsymbol{\alpha}$ are both sets of uncertain variables, where $\boldsymbol{\delta} = (\delta_i)_{i \in \Lambda}^{n_c \times 1}$ with $i \in \Lambda$ and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$. The sub-objective F_1 representing the total investment cost of STATCOMs to be installed is given by

$$F_1(\mathbf{x}) = \sum_{i \in \Lambda} (x_i C(x_i)) \quad (8)$$

with the cost function $C(x_i)$ defining the relationship between the cost per unit of STATCOMs and rated capacity of STATCOMs installed at bus i . Also, $C(x_i)$ used in this work is a piecewise function given as

$$C(x_i) = \begin{cases} c_1 & 0 < x_i \leq x_{p1} \\ c_2 & x_{p1} < x_i \leq x_{p2} \\ \dots \end{cases} \quad (9)$$

with c_1 and c_2 being constants, and x_{p1} and x_{p2} denoting piecewise points [24]. Other forms of cost function can be employed by the planner according to specific circumstances.

The sub-objectives F_2 and F_3 are the expectancy measure and dispersion measure of the short-term voltage stability level, respectively, representing the mean performance and robustness of planning solutions w.r.t. the short-term voltage stability level under uncertainties. Before giving formulations of F_2 and F_3 , a function $\psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha})$ denoting the short-term voltage stability level of power grids, with $\mathbf{x} - \boldsymbol{\delta}$ being the vector of the actual capacity of STATCOMs when contingencies occur under uncertainties $\boldsymbol{\alpha}$ is first introduced. Also, $\psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha})$ is calculated by (10) according to post-disturbance voltage trajectories obtained by time-domain simulations:

$$\psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha}) = \sum_{k \in \Theta} \left\{ \frac{p_k \left(\sum_{i \in \mathbf{I}} \sum_{t \in \Gamma} \Delta t \left([\underline{V}_{i,t} - V_{i,t}^k]^+ + [V_{i,t}^k - \bar{V}_{i,t}]^+ \right) \right)}{n_b T} \right\}, \quad (10)$$

where \mathbf{I} and Γ are sets of buses of power grids and simulation points in T , the time span of short-term voltage stability from the contingency cleared, respectively. Θ and p_k are the same as that in (2), n_b denotes the number of buses in \mathbf{I} , Δt is the time step of time-domain simulations, $V_{i,t}^k$ represents voltage amplitude of bus i at time t after the contingency k , and $[\cdot]^+ = \max(0, \cdot)$. Variables \underline{V}_i and \bar{V}_i , together depicting boundaries which the more voltage trajectories overstep, the more short-term voltage stability level is reduced, are given by

$$\begin{cases} \underline{V}_t = \frac{V_{st}(\exp(t/t_{end})t/t_{end})^\beta}{\exp(\beta)}, & \forall t \in \Gamma \\ \bar{V}_t = 2 - \underline{V}_t, \end{cases} \quad (11)$$

with t_{end} denoting the end instant in Γ and β being a decay coefficient used to calibrate the decaying of the boundaries towards the steady-state value V_{st} .

Given \mathbf{x} , $\psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha})$ is a function of only uncertainties $\boldsymbol{\delta}$ and $\boldsymbol{\alpha}$, and we regard it as a response $R = \psi(\boldsymbol{\delta}, \boldsymbol{\alpha})$. Taking $n_c = 2$ for example, the relationship between input uncertainties and response R is illustrated in Fig. 1. Due to the mixture of aleatory uncertainties $\boldsymbol{\delta}$ and epistemic uncertainties $\boldsymbol{\alpha}$, the uncertainty of R is in the form of a family of probability distributions. Also, each probability distribution is the response to input uncertainties $\boldsymbol{\alpha}$ given fixed values of input uncertainties $\boldsymbol{\delta}$.

The sub-objective F_2 , an expectancy measure, is defined by the average mean value of R as follows:

$$F_2(\mathbf{x}, \boldsymbol{\delta}, \boldsymbol{\alpha}) = \frac{1}{2} (\mu_R^{\max} + \mu_R^{\min}), \quad (12)$$

$$f_{\alpha_3}(\alpha_3 | \mu_{31}, \mu_{32}, \delta_3, a_3, b_3) = \begin{cases} \frac{1}{\sigma_3} \phi\left(\frac{\alpha_3 - \mu_{31}}{\sigma_3}\right) + \frac{1}{\sigma_3} \phi\left(\frac{\alpha_3 - \mu_{32}}{\sigma_3}\right) \\ \Phi\left(\frac{b_3 - \mu_{31}}{\sigma_3}\right) - \Phi\left(\frac{a_3 - \mu_{31}}{\sigma_3}\right) + \Phi\left(\frac{b_3 - \mu_{32}}{\sigma_3}\right) - \Phi\left(\frac{a_3 - \mu_{32}}{\sigma_3}\right), & a_3 \leq \alpha_3 \leq b_3, \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

where $\mu_R^{\max} = \max \{\mu_{R_1}, \mu_{R_2}, \dots, \mu_{R_{2^{n_c}}}\}$ and $\mu_R^{\min} = \min \{\mu_{R_1}, \mu_{R_2}, \dots, \mu_{R_{2^{n_c}}}\}$ with μ_{R_i} denoting the mean value of the probability distribution of R . Also, μ_{R_i} can be written as

$$\mu_{R_i} = E[\psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha}) | \mathbf{x}, \boldsymbol{\delta}] = \int_{\Omega} \psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha}) p(\boldsymbol{\alpha}) d\boldsymbol{\alpha}, \quad (13)$$

where Ω stands for the support region of $\boldsymbol{\alpha}$, and $p(\boldsymbol{\alpha})$ represents the joint PDF of $\boldsymbol{\alpha}$, given by

$$p(\boldsymbol{\alpha}) = f_{\alpha}(a_1 | \mu_1, \sigma_1, a_1, b_1) \times f_{\alpha}(a_2 | \mu_2, \sigma_2, a_2, b_2) \times f_{\alpha}(a_3 | \mu_3, \sigma_3, a_3, b_3). \quad (14)$$

The sub-objective F_3 , a dispersion measure, is defined by the sum of two items: the average standard deviation of R ($\bar{\sigma}_R$) used as a dispersion measure for aleatory uncertainties $\boldsymbol{\alpha}$ and the standard deviation difference of R ($\Delta\sigma_R$) used as a dispersion measure for epistemic uncertainties $\boldsymbol{\delta}$. Also, F_3 is given by

$$F_3(\mathbf{x}, \boldsymbol{\delta}, \boldsymbol{\alpha}) = \bar{\sigma}_R + \Delta\sigma_R = \frac{1}{2}(\sigma_R^{\max} + \sigma_R^{\min}) + (\sigma_R^{\max} - \sigma_R^{\min}), \quad (15)$$

where $\sigma_R^{\max} = \max \{\sigma_{R_1}, \sigma_{R_2}, \dots, \sigma_{R_{2^{n_c}}}\}$ and $\sigma_R^{\min} = \min \{\sigma_{R_1}, \sigma_{R_2}, \dots, \sigma_{R_{2^{n_c}}}\}$ with σ_{R_i} denoting the standard deviation of a probability distribution of R . Also, σ_{R_i} can be written as

$$\sigma_{R_i} = \sqrt{E[(\psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha}) - \mu_{R_i})^2 | \mathbf{x}, \boldsymbol{\delta}]} = \sqrt{E[(\psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha}))^2 | \mathbf{x}, \boldsymbol{\delta}] - \mu_{R_i}^2}, \quad (16)$$

where

$$E[(\psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha}))^2 | \mathbf{x}, \boldsymbol{\delta}] = \int_{\Omega} (\psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha}))^2 p(\boldsymbol{\alpha}) d\boldsymbol{\alpha}. \quad (17)$$

Here, decision makers will prefer smaller values of F_2 and F_3 since they correspond to planning solutions with better mean value behaviour and stronger robustness with respect to the short-term voltage stability level, respectively.

3.2 Constraints of the proposed optimisation model

The constraints of the optimisation model consist of two categories: constraints for power grids and constraints for decision variables. The first category of constraints consists of DAE constraints as given in (18), and pre-disturbance steady-state operation constraints including voltage amplitude constraints, generators output constraints, etc., which can be satisfied by choosing the appropriate operating condition

$$\begin{cases} \dot{\mathbf{u}} = \mathbf{f}(\mathbf{x} - \boldsymbol{\delta}, \mathbf{u}, \mathbf{y}), \\ \mathbf{0} = \mathbf{g}(\mathbf{x} - \boldsymbol{\delta}, \mathbf{u}, \mathbf{y}), \end{cases} \quad (18)$$

where \mathbf{u} and \mathbf{y} are vectors of state variables and algebraic variables, respectively.

Also, constraints for \mathbf{x} are as follows:

$$\begin{cases} 0 \leq x_i \leq x_{\max}, & \forall i \in \Lambda, \\ x_i = z \Delta x, & \forall i \in \Lambda, \exists z \in \mathbb{Z}_0^+, \end{cases} \quad (19)$$

where x_{\max} and Δx are the maximum rated capacity and minimum capacity interval of STATCOMs, respectively.

4 Methodology for solving the proposed multi-objective robust dynamic VAR planning model

Involvement of uncertainties in the proposed optimisation model increases the computational expense of the solving process substantially, compared with the computational work of classical deterministic planning model. Also, it is important to develop and implement a computationally efficient methodology for solving the proposed multi-objective optimisation model.

Surrogate modelling technologies have been widely applied to optimisation problems in many fields such as aerodynamic design to cope with unaffordable computing burdens caused by computational simulations (e.g. computational fluid dynamics and finite element analysis) [25]. Also, put simply, the surrogate modelling technology mimics the behaviour of the original simulation model accurately while being computationally cheaper to evaluate, by properly constructing simpler approximate models.

Motivated by this, in this section, a methodology using multi-objective evolutionary algorithms (MOEAs) combined with surrogate modelling technologies is proposed for solving the optimisation model. Other issues in the proposed methodology, including model reduction before optimising, uncertainty quantification, and training data sampling, are also carefully addressed in this section.

4.1 Model reduction before optimising

In analogy to solving the classical multi-objective dynamic VAR planning model, it is indispensable to reduce the proposed optimisation model by selecting severe contingencies and potential location buses. In this work, the clustering based selection of severe contingencies and global sensitivity analysis based selection of potential location buses proposed in our previous work [15] are employed. Also, here, a brief overview of them is given and more details can be found in [15].

In the clustering-based selection of severe contingencies, firstly, $\forall k \in \Theta$, calculate $\psi|_{\Theta=k, p_k=1}$ by time-domain simulations. Also, then contingencies in Θ , causing short-term voltage instability, are identified according to time-domain simulation results and added to the set of severe contingencies Θ_s . Following that, other contingencies in Θ causing no instability are clustered into a set of two clusters by K-means using $\psi|_{\Theta=k, p_k=1}$ as the feature of contingencies. Finally, the cluster with more severe contingencies is further added to Θ_s to obtain the final Θ_s . It should be mentioned that the contingencies causing short-term voltage instability here refer to those causing instability and in which voltage is the domain factor since we only focus on the improvement of short-term voltage stability level by installing STATCOMs in this study.

Also, in the global sensitivity analysis based selection of potential location buses, $\forall i \in \Lambda$, the global sensitivity measure δ_i^A that implies the ability of buses to improve short-term voltage stability level by being installed STATCOMs is calculated. Also, the first n_c initial candidate buses with the largest values of global sensitivity measure are assigned to the set of potential location buses Λ_c .

4.2 Uncertainty quantification method

Since in the uncertainty quantification in terms of $\boldsymbol{\alpha}$, i.e. the estimation of μ_{R_i} and σ_{R_i} , the expensive function ψ will be replaced by its surrogate model $\tilde{\psi}$ that will be given in the later section, the uncertainty quantification method used is not demanding for efficiency in terms of computation speed. Also, thus in our work, we use a sampling-based method with Latin hypercube sampling for the uncertainty quantification.

The uncertainty quantification method contains two steps. The first step is using Latin hypercube sampling to generate a set of samples of random variables $\boldsymbol{\alpha}$ obeying the joint distribution $p(\boldsymbol{\alpha})$, denoted by $\chi = \{\boldsymbol{\alpha}_i, i = 1, 2, \dots, n_s\}$ with $\boldsymbol{\alpha}_i = (\alpha_{1,i}, \alpha_{2,i}, \alpha_{3,i})$ and n_s being the number of samples. In this step, the CDF of each variable in $\boldsymbol{\alpha}$ is firstly evenly partitioned into n_s regions and one random sample is generated in each region. Also, then samples of each variable are randomly combined into three-dimensional pairs, i.e.

$(\alpha_{1,i}, \alpha_{2,i}, \alpha_{3,i})$. Also, in the second step, μ_{R_i} and σ_{R_i} are estimated from the samples χ . The equations are as follows:

$$\mu_{R_i} \cong \frac{1}{n_s} \sum_{i=1}^{n_s} \psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha}_i), \quad (20)$$

$$\sigma_{R_i} \cong \sqrt{\frac{1}{n_s} \sum_{i=1}^{n_s} (\psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha}_i))^2 - \mu_{R_i}^2}. \quad (21)$$

4.3 Multi-objective optimisation algorithm

MOEAs have been successfully used to solve the classical multi-objective dynamic VAR planning model [11, 15, 16]. In our work, a highly reliable and efficient MOEA, called ϵ -NSGAII, is applied to solve the proposed optimisation model. This paper assumes a basic prior knowledge of MOEAs and only focuses on the use of ϵ -NSGAII. The interested readers can refer to [26, 27] for additional details.

Three parameters, i.e. ϵ_i , N_0 and $\Delta\%$, are user-provided when using ϵ -NSGAII. ϵ_i defines the users' precision goal of the sub-objective F_i . Smaller ϵ_i values result in higher precision and ultimately more solutions. N_0 denotes the initial population size from which the population size is automatically adapted during the search process. Also, $\Delta\%$ defines the termination condition. The search is automatically terminated if the number and quality of optimal solutions have not increased above $\Delta\%$ across two successive runs.

4.4 Extreme learning machine (ELM)-based surrogate modelling

One of the most important steps of the proposed solving methodology is to build the surrogate model for the expensive function $\psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha})$. In this section, we propose an ELM-based surrogate modelling method with adaptive training data sampling.

4.4.1 Surrogate model construction: In our work, ELM, a powerful emergent machine learning technique, is used for surrogate model construction [28]. Compared with other analogous techniques such as support vector regression, ELM can solve any regression problem with the desired accuracy and faster training time as well [29].

An ELM is a single-layer feed-forward network with a fast training method. Consider a set of N distinct training data (\mathbf{x}'_i, t_i) , $i \in [1, N] \cap \mathbb{Z}$ with $\mathbf{x}'_i = (\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha})$ and $t_i = \psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha})$. Then the output of the hidden layer with L hidden neurons is

$$\mathbf{H} = \begin{bmatrix} \varphi(\mathbf{w}_1 \mathbf{x}'_1 + b'_1) & \dots & \varphi(\mathbf{w}_L \mathbf{x}'_1 + b'_L) \\ \vdots & \ddots & \vdots \\ \varphi(\mathbf{w}_1 \mathbf{x}'_N + b'_1) & \dots & \varphi(\mathbf{w}_L \mathbf{x}'_N + b'_L) \end{bmatrix}, \quad (22)$$

where φ is the activation function (the sigmoid function is used in our work), $\mathbf{W} = (\mathbf{w}_1 \dots \mathbf{w}_L)$ is the input weight and $\mathbf{b} = (b'_1 \dots b'_L)$ is the bias. In the ELM model, \mathbf{W} and \mathbf{b} are set randomly and never adjusted. Also, then with $\boldsymbol{\beta} = (\beta_1 \dots \beta_L)^T$ representing the output weight, the target output $\mathbf{t} = (t_1 \dots t_N)$ and the estimated output $\mathbf{y} = (y_1 \dots y_N)$ are given by

$$\mathbf{y} = \mathbf{t} + \boldsymbol{\epsilon} = \mathbf{H}\boldsymbol{\beta}, \quad (23)$$

where $\boldsymbol{\epsilon}$ is noise.

Generally, an ELM model is over-determined with $N > L$. Also, a unique solution for (23) can be given by a minimum L_2 norm of the training error using the Moore–Penrose generalised inverse of the matrix \mathbf{H} , denoted as \mathbf{H}^\dagger . Thus, we have

$$\boldsymbol{\beta} = \mathbf{H}^\dagger \mathbf{t}, \quad \mathbf{H}^\dagger = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T. \quad (24)$$

Also, finally, the surrogate model of $\psi(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha})$ can be written as

$$\tilde{\psi}(\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha}) = \sum_{j=1}^L \beta_j \varphi(\mathbf{w}_j \mathbf{x}' + b'_j), \quad (25)$$

where $\mathbf{x}' = (\mathbf{x} - \boldsymbol{\delta}, \boldsymbol{\alpha})$.

4.4.2 ELM structure selection: In the surrogate model, the number of hidden neurons that determine the ELM structure can greatly affect the model performance, i.e. the model prediction error. Thus in the construction process, a K -fold cross-validation strategy coupled with the mean squared error metric is applied to measure the prediction error and guide the ELM structure selection.

Given the number of hidden neurons, the K -fold cross-validation strategy firstly divides all the training data into K groups of data, called folds. For the k th fold, the surrogate model is trained using the other $k-1$ folds, and the prediction error of the trained model is estimated using the root mean squared error with the k th fold

$$\text{RMSE}(\tilde{\psi}^{-k}) = \sqrt{\frac{1}{|X_k|} \sum_{(\mathbf{x}'_i, t_i) \in X_k} (t_i - \tilde{\psi}^{-k}(\mathbf{x}'_i - \boldsymbol{\delta}_i, \boldsymbol{\alpha}_i))^2}. \quad (26)$$

Here, $\tilde{\psi}^{-k}$ denotes the surrogate model trained after removing the k th fold of training data. Also, X_k represents the set of training data in the k th fold and $(\mathbf{x}'_i - \boldsymbol{\delta}_i, \boldsymbol{\alpha}_i) = \mathbf{x}'_i$. Then this process is repeated for $k = 1, 2, \dots, K$, and an estimation for the prediction error of the surrogate model with the given ELM structure can be expressed by the mean of all $\text{RMSE}(\tilde{\psi}^{-k})$

$$\text{MRMSE} = \frac{1}{K} \sum_{k=1}^K \text{RMSE}(\tilde{\psi}^{-k}). \quad (27)$$

Also, smaller MRMSE values indicate smaller prediction errors.

The number of hidden neurons is gradually increased by an interval ΔL , and respective MRMSE is calculated with the average results of repetitive random trials of simulations for each fixed ELM structure. Also, finally, the nearly optimal ELM structure is determined by the number of hidden neurons followed that the mean value of MRMSE first begins to increase.

4.4.3 Adaptive training data sampling: In the surrogate model construction, proper training data sampling is crucial for achieving an adequate accuracy of the surrogate model that allows us to guide the optimisation process towards optimal points of the original model. However, the classical sampling methods, such as the space-filling method and Latin hypercube sampling method, need intensively sampling the input space for a satisfactory accuracy. This can eliminate the merits of the application of surrogate modelling since training data is obtained by performing expensive time-domain simulations on the selected points. In this section, a novel adaptive training data sampling method is proposed and this method, using acceptable computational efforts, allows a good coverage of the input space while providing enough accuracy in regions where optimal points are potentially located.

Initial training data is first obtained through a crossed design for input variables $\mathbf{x} - \boldsymbol{\delta}$ related to decision variables and uncertainty variables $\boldsymbol{\alpha}$; i.e. n_d combinations of variables $\mathbf{x} - \boldsymbol{\delta}$ are combined with n_u combinations of variables $\boldsymbol{\alpha}$. These n_d combinations are sampled through the Latin hypercube sampling method described in Section 4.2, assuming that each variable in $\mathbf{x} - \boldsymbol{\delta}$ has a continuous uniform distribution on $[0, x_{\max}]$. Also, the n_u combinations are space-filling. These $n_d \times n_u$ combinations, i.e. points of input, are then simulated to obtain the output part of the initial training data. Note that n_u should be relatively small to reduce the computational effort while n_u should be large enough since the sampling of $\boldsymbol{\alpha}$ will not be updated later. Also, an initial Pareto set of the proposed multi-objective optimisation model can

be obtained by executing ϵ -NSGAI I with the surrogate model constructed using initial training data.

Then initial training data is updated by adding extra ones to improve both the coverage of the input space and accuracy in regions where initial Pareto-optimal solutions are located. More specifically, we add training points that are far away from the existing ones or near regions where initial Pareto-optimal solutions are located but prediction accuracy is insufficient. Following that, a more accurate Pareto set can be potentially obtained by executing ϵ -NSGAI I with the surrogate model constructed using the updated training data. Also, these two processes, namely updating training data and executing ϵ -NSGAI I , are repeated until a certain termination condition is satisfied.

Mathematically, before the $(n+1)$ th repeat, the up-to-date Pareto set and the set of combinations of variables $\mathbf{x} - \boldsymbol{\delta}$ in the up-to-date training data $X^{(n)}$ are denoted by $X_p^{(n)}$ and $X_\delta^{(n)}$, respectively. Also, n_{da} combinations of variables $\mathbf{x} - \boldsymbol{\delta}$ are assumed to be added in one repeat, and among which $n_{dc}^{(n)}$ combinations aim to improve the coverage of the input space. In addition, two auxiliary functions g_a and h_a are introduced as follows:

$$g_a(\mathbf{x}_\delta, X_\delta^{(n)}) = \min_{\mathbf{x}_{\delta,i} \in X_\delta^{(n)}} d(\mathbf{x}_\delta, \mathbf{x}_{\delta,i}), \quad (28)$$

$$h_a(\mathbf{x}_{\delta,i}, X^{(n)}) = \sqrt{\frac{\sum_{j=1}^{n_u} [\tilde{\psi}^{-i}(\mathbf{x}_{\delta,i}, \boldsymbol{\alpha}_j) - \psi(\mathbf{x}_{\delta,i}, \boldsymbol{\alpha}_j)]^2}{n_u}}. \quad (29)$$

In (28), \mathbf{x}_δ denotes a generic combination of variables $\mathbf{x} - \boldsymbol{\delta}$, and $d(\mathbf{x}_\delta, \mathbf{x}_{\delta,i})$ represents the Euclidean distance between \mathbf{x}_δ and $\mathbf{x}_{\delta,i}$. In (29), $\mathbf{x}_{\delta,i} \in X_\delta^{(n)}$ and $\boldsymbol{\alpha}_j$ is one of the n_u combinations of $\boldsymbol{\alpha}$ obtained before. The function $\tilde{\psi}^{-i}$ represents the surrogate model trained after removing the training data that $\mathbf{x} - \boldsymbol{\alpha} = \mathbf{x}_{\delta,i}$ from $X^{(n)}$, and the value of $\psi(\mathbf{x}_{\delta,i}, \boldsymbol{\alpha}_j)$ can be obtained from $X^{(n)}$. Larger h_a values indicate higher prediction errors in the region around $\mathbf{x}_{\delta,i}$ with the surrogate model trained using $X^{(n)}$.

With the auxiliary functions g_a and h_a , the updating process of training data is given by Algorithm 1. In Algorithm 1, the $n_{da} \times n_u$ combinations are analogous to the $n_d \times n_u$ combinations in the initialisation phase and $X^{(n+1)}$ denotes training data after being updated. Moreover, the random seed used for Latin hypercube sampling here should be changed for different n in order to avoid duplication.

Algorithm 1: Updating algorithm of training data

Input: $X^{(n)}, X_p^{(n)}; n_{da}; n_{dc}^{(n)}$

- 1: $X_h \leftarrow$ a huge set of combinations of variables $\mathbf{x} - \boldsymbol{\delta}$ sampled by Latin hypercube sampling
- 2: $k \leftarrow 0, \Delta X_\delta \leftarrow \{\}, H_a \leftarrow \{\}$
- 3: **repeat**
- 4: $\Delta X_\delta \leftarrow \Delta X_\delta + \arg \max_{\mathbf{x}_{\delta,i} \in X_h} g_a(\mathbf{x}_{\delta,i}, X_\delta^{(n)} \cup \Delta X_\delta)$
- 5: $k \leftarrow k + 1$
- 6: **until** $k = n_{dc}^{(n)}$
- 7: **for all** $\mathbf{x}_i \in X_p^{(n)}$ **do**
- 8: $\mathbf{x}_{\delta,i}^* \leftarrow \arg \min_{\mathbf{x}_{\delta,j} \in X_\delta^{(n)}} d(\mathbf{x}_i, \mathbf{x}_{\delta,j})$
- 9: $h_{a,i} \leftarrow h_a(\mathbf{x}_{\delta,i}^*, X^{(n)})$
- 10: $H_a \leftarrow H_a + h_{a,i}$
- 11: **end for**
- 12: $X_l \leftarrow$ the set of $\mathbf{x}_i \in X_p^{(n)}$ with the first $n_{da} - n_{dc}^{(n)}$ largest $h_{a,i}$ in H_a
- 13: **for all** $\mathbf{x}_i \in X_l$ **do**
- 14: $\Delta X_\delta \leftarrow \Delta X_\delta + \arg \min_{\mathbf{x}_{\delta,i} \in X_h} d(\mathbf{x}_i, \mathbf{x}_{\delta,i})$
- 15: **end for**
- 16: Simulate the $n_{da} \times n_u$ combinations to obtain extra training data ΔX

17: $X^{(n+1)} \leftarrow X^{(n)} \cup \Delta X$

Output: $X^{(n+1)}$

Another important issue in Algorithm 1 is the trade-off between computational efforts used for improving the coverage of the input space and those used for improving the accuracy in regions where up-to-date Pareto-optimal solutions are located. When the available training data lacks coverage of the input space, the prediction error is supposedly higher away from the available data. Also, if extra training data is sampled from the blank area of the input space, the resulting Pareto set, assumed to be $X_p^{(n+1)}$, is apt to differ widely from the previous one $X_p^{(n)}$. Thus the dissimilarity between Pareto sets is used to guide this trade-off by adaptively adjusting the value of $n_{dc}^{(n)}$ during the repeat process.

The Jaccard distance is first introduced to measure the dissimilarity between Pareto sets obtained from two successive repeats and is given by [30]

$$d_J(X_p^{(n+1)}, X_p^{(n)}) = \frac{|X_p^{(n+1)} \cup X_p^{(n)}| - |X_p^{(n+1)} \cap X_p^{(n)}|}{|X_p^{(n+1)} \cup X_p^{(n)}|}. \quad (30)$$

Here, the range of d_J is $[0, 1]$ and larger d_J values indicate a greater degree of dissimilarity between $X_p^{(n+1)}$ and $X_p^{(n)}$. Also, then the adaptive adjustment of $n_{dc}^{(n)}$ is given by

$$n_{dc}^{(n)} = \begin{cases} n_{da} & n = 0 \\ \lceil n_{da} d_J(X_p^{(n)}, X_p^{(n-1)}) \rceil & n > 0 \end{cases} \quad (31)$$

with $\lceil \cdot \rceil$ denoting the ceiling function. The $n_{dc}^{(n)}$ is initialised to n_{da} , namely that all computational efforts are assigned to improve the coverage of the input space. Also, when $X_p^{(n)}$ is identical to $X_p^{(n-1)}$, computational efforts will be entirely used to improve the accuracy around Pareto-optimal solutions.

4.5 Overall flow

The overall flow of the proposed methodology for solving the multi-objective robust dynamic VAR planning model is given in Fig. 2. Note that the ELM structure selection is only conducted when constructing the initial ELM surrogate model, i.e. $n = 0$, and the obtained nearly optimal ELM structure is also used for the later construction of ELM models. In addition to the aforementioned processes given before this section, the stop condition in Fig. 2 is formulated based on the distance between Pareto fronts obtained from two successive repeats. As the repeat process progresses, the Pareto set and Pareto front first drastically move in the decision variable space and objective space, respectively. The Pareto set will then tend to be fixed after training data covers the input space thoroughly. Also, following that, F_2 and F_3 of the Pareto front will approximate to their true values as extra training data continuously being added around the Pareto set. Theoretically, the absolute value of the maximum difference between F_2 (or F_3) values obtained from two successive repeats will monotonically decrease to zero since the extra training data is always added to regions where prediction errors are the highest. Thus the repeat process can stop if the Pareto set is not changed and the maximum difference between F_2 (or F_3) values obtained from two successive repeats is less than a given small value. Mathematically, the stop condition can be written as

$$d_J(X_p^{(n)}, X_p^{(n-1)}) = 0 \wedge \gamma^{(n)} < \gamma \quad (32)$$

with

$$\gamma^{(n)} = \max \{ \max \{ |F_{2,i}^{(n)} - F_{2,i}^{(n-1)}|, |F_{3,i}^{(n)} - F_{3,i}^{(n-1)}| \} : i = 1, 2, \dots, |X_p^{(n)}| \}, \quad (33)$$

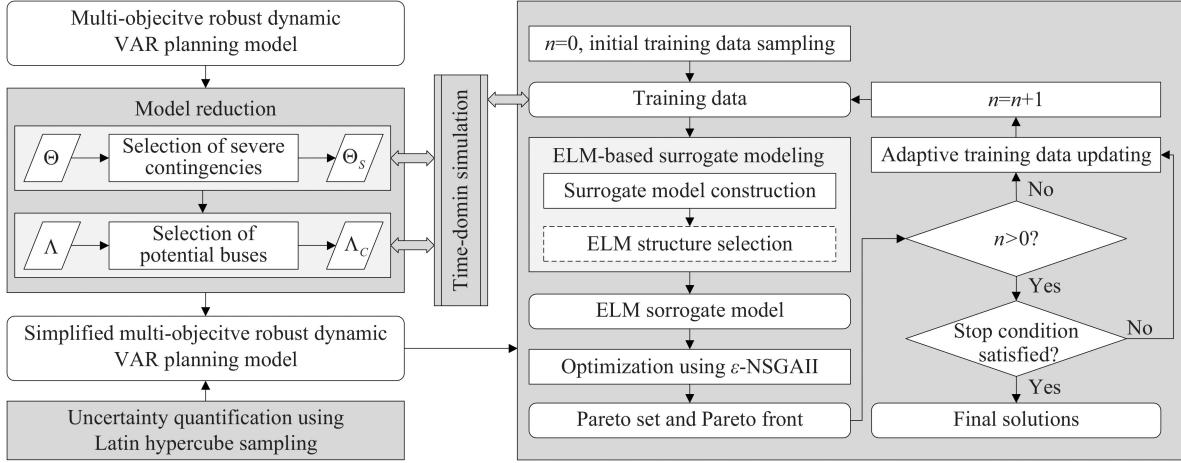


Fig. 2 Overall flowchart of the proposed methodology for solving the multi-objective robust dynamic VAR planning model

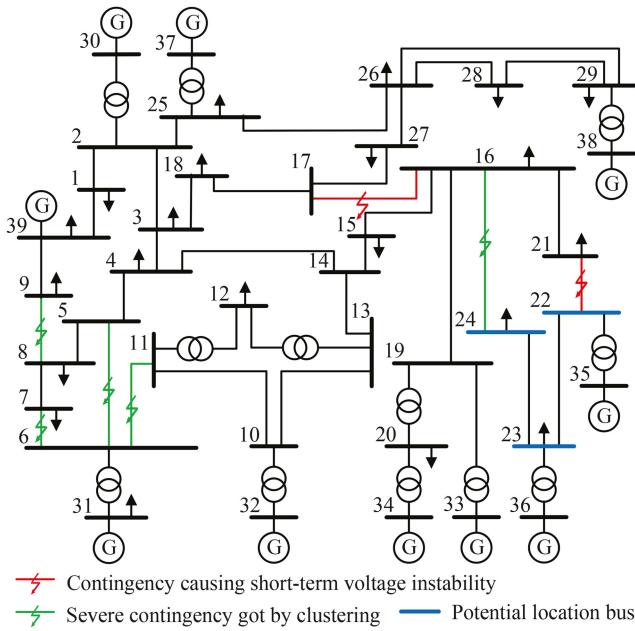


Fig. 3 One-line diagram of the IEEE 39-bus system

Table 1 Parameters of uncertainty variables

Uncertainty variable	Parameter
a_1	$\mu_1 = 6254.23 \text{ MW}$, $\sigma_1 = 125.0846 \text{ MW}$, $a_1 = 5941.5185 \text{ MW}$, $b_1 = 6566.9415 \text{ MW}$
a_2	$\mu_2 = 50\%$, $\sigma_2 = 2\%$, $a_2 = 45\%$, $b_2 = 55\%$
a_3	$\mu_{31} = 0.09 \text{ s}$, $\mu_{32} = 0.1 \text{ s}$, $\sigma_3 = 0.0048 \text{ s}$, $a_3 = 0.074 \text{ s}$, $b_3 = 0.116 \text{ s}$
$\delta_i, \forall i \in \Lambda$	$\underline{\delta}_i = 0$, $\bar{\delta}_i = 0.2x_i$

Table 2 Cost function of STATCOMs

$x_i, \text{ MVar}$	$0 < x_i \leq 100$	$100 < x_i \leq 200$	$200 < x_i \leq 250$
$C(x_i) (\$/\text{million/MVar})$	0.13	0.115	0.11

where $F_{2,i}^{(n)}$ denotes values of the objective F_2 of the i th Pareto-optimal solution in $X_p^{(n)}$, and $F_{2,i}^{(n-1)}$, $F_{3,i}^{(n)}$ and $F_{3,i}^{(n-1)}$ are analogous. The term γ represents a given small value.

5 Case studies

The proposed multi-objective robust dynamic VAR planning approach is tested on the IEEE 39-bus system [31]. The one-line

diagram of the test system is shown in Fig. 3. The PSD-BPA (Power System Department Bonneville Power Administration, a power system analysis and simulation software) is used to perform time-domain simulations, and ϵ -NSGAII is implemented by the MOEA framework [32], an open source Java framework for multi-objective optimisation. Other computing is implemented in PYTHON. An interface between PYTHON and PSD-BPA, and the MOEA framework, developed in our previous work is also applied [15]. A standard 64-Bit personal computer system with Intel(R) Core(TM) i7-6700 3.40 GHz CPU, and 8 GB RAM is used.

5.1 Parameter settings

Parameters of uncertainty variables are given in Table 1. The distribution of loads under different load levels is assumed to be the same as that in the original data. In addition, the load model consists of α_2 dynamic load and $(1 - \alpha_2)$ static load. Parameters of dynamic load are taken from that of small industrial inductors in [33], and static load consists of 40% constant impedance load, 30% constant current load and 30% constant power load.

The maximum rated capacity x_{\max} and minimum capacity interval Δx of STATCOMs are assumed to be 250 and 5 MVar, respectively. In practice, x_{\max} and Δx depend on the manufacturing level and actual conditions. The cost function of STATCOMs is shown in Table 2 [24]. The dynamic model of STATCOMs can be found in [34] and their parameters are obtained from practical engineering.

All three-phase short circuit faults of branches causing no isolation are taken as credible contingencies. Also, all faults are assumed to occur in the middle of branches and can be cleared after a time of α_3 by disconnecting two sides' breakers. It is also assumed that $\forall k \in \Theta$, $p_k = 1/|\Theta|$. Moreover, all non-generator buses in the power grid are taken as initial candidate buses to install STATCOMs from which potential location buses are further selected. Also, for real power grids, bus location and substation physical size should be also taken into account.

For the calculation of ψ , $\beta = 0.024$ and $V_{st} = 0.9 \text{ p.u.}$ are used [13]. Also, the time span T is set to 5 s. For time domain simulations, the simulation step Δt is set to 0.01 s, and the total simulation time for one fault is set to 5.5 s.

In the uncertainty quantification in terms of α , 1500 samples are generated using Latin hypercube sampling [35]. In ϵ -NSGAII, parameter ϵ_i is set to $\Delta x \times \min C(x_i) = 5 \times 0.11 = 0.55$ for $i = 1$, and $0.01F_{i,\max}$ for $i = \{2, 3\}$, with $F_{i,\max}$ denoting the maximum value of F_i , $F_{2,\max} = F_2(\mathbf{0}, \boldsymbol{\delta}, \boldsymbol{\alpha})$ and $F_{3,\max} = F_3(\mathbf{0}, \boldsymbol{\delta}, \boldsymbol{\alpha})$ which can be obtained with the surrogate model constructed using initial training data. In addition, $\Delta\% = 10\%$ and $N_0 = 10$ are used.

In the ELM structure selection, 10-fold cross-validation is used and ΔL is set to 500. Also, 10 random trials of simulations for each fixed ELM structure are conducted for calculating MRMSE. For training data sampling, n_d is set to 20 times the dimension of input

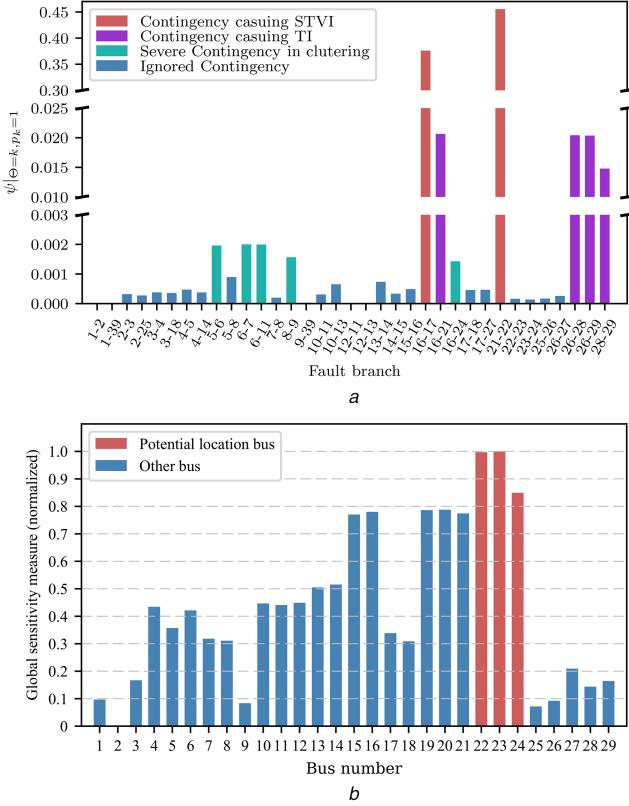


Fig. 4 Results of model reduction before optimising. STVI and TI are abbreviations of short-term voltage instability and transient instability, respectively

(a) Selection of severe contingencies, (b) Selection of potential location buses

variables $x - \delta$, n_u is set to 5^3 , and n_{da} is set to 5. In addition, $\gamma = 0.1 \min\{\epsilon_2, \epsilon_3\}$ is used for termination of the repeat process.

5.2 Results of model reduction before optimising

Fig. 4 shows results of model reduction before optimising, including selection results of severe contingencies in Fig. 4a and those of potential location buses in Fig. 4b. In Fig. 4a, contingencies occurring in branches 16–17, 21–22, 16–21, 26–28, 26–29 and 28–29 are with large values of $\psi|_{\Theta=k, p_k=1}$ and all cause instability according to their post-disturbance voltage curves and oscillation curves, which can be found in our previous work [15]. Also, in contingencies occurring in branches 16–17 and 21–22, bus voltages continuously drop, leading further loss of synchronism of machines. Thus voltages are the domain factor for these two contingencies. However, rotor angles are the domain factor for contingencies occurring in branches 16–21, 26–28, 26–29 and 28–29 since the loss of synchronism of two groups of machines causes rapid drop and oscillation in voltages in these contingencies. Moreover, other contingencies with small values of $\psi|_{\Theta=k, p_k=1}$ and causing no obvious instability are further clustered into two clusters. Also, contingencies coloured green are the cluster with more severe ones. Finally, the set Θ_S consists of contingencies occurring in branches 16–17, 21–22, 5–6, 6–7, 6–11, 8–9 and 16–24. Also, in latter calculations, $\forall k \in \Theta_S, p_k = 1/7$.

From Fig. 4b, it can be seen that values of the global sensitivity measure vary among different buses, implying that these buses have different abilities to improve the short-term voltage stability level by installing STATCOMs. Also, in this test, with $n'_c = 3$, we have $\Lambda_c = \{22, 23, 24\}$ (red buses in Fig. 4b).

5.3 Results of optimisation

After collecting the initial training data with $20 \times 3 \times 5^3 = 7500$ samples, the ELM structure selection is conducted for the initial surrogate modelling. Runtime results of MRMSE versus the

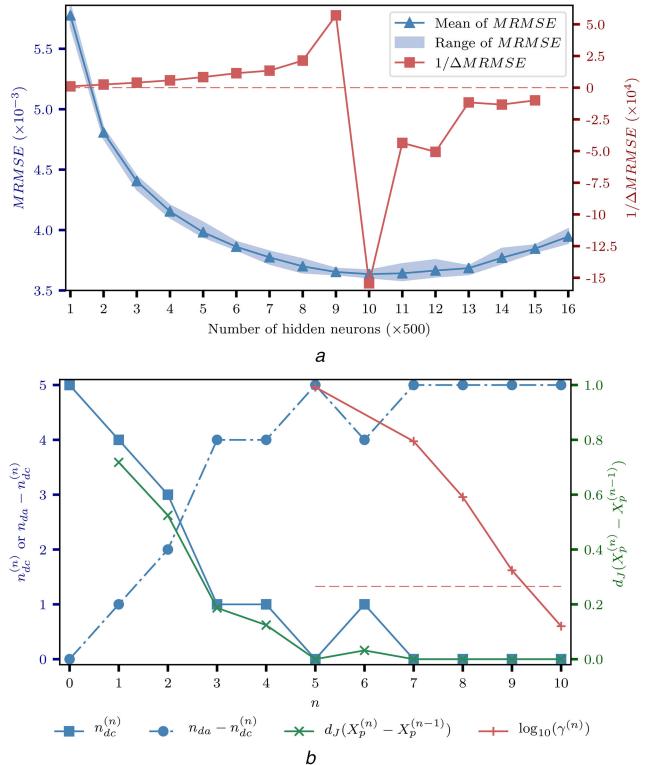


Fig. 5 Change plots of MRMSE versus the number of hidden neurons and key variables in the repeat process

(a) Change plot of MRMSE versus the number of hidden neurons for ten random trials of simulations. $\Delta MRMSE$ denotes the difference between the mean of MRMSE with the current number of hidden neurons and that with the next tested number of hidden neurons. Also, the reciprocal of $\Delta MRMSE$ is used here to make small values of $\Delta MRMSE$ more visible, (b) Change plot of key variables including $n_{dc}^{(n)}$, $n_{da} - n_{dc}^{(n)}$, $d_J(X_p^{(n)} - X_p^{(n-1)})$ and $\gamma^{(n)}$ in the repeat process. Also, $\gamma^{(n)}$ is log-transformed to make the end point more visible since $\gamma^{(n)}$ decreases nearly exponentially

number of hidden neurons are represented in Fig. 5a. It can be seen that with the growth of the number of hidden neurons, the mean of MRMSE is reduced followed by a gradual increase. Also, $\Delta MRMSE$ is first small than 0 when the number of hidden neurons goes up to 5000. In other words, the number of hidden neurons followed that values of MRMSE first begin to increase is 5000. Thus the nearly optimal ELM structure is with 5000 hidden neurons. Note that the ELM structure selection process can be terminated when the number of hidden neurons increases to 5500 but the process is continued here to make the changing plot more distinct.

Also, using the initial trained ELM model with the nearly optimal structure, parameters ϵ_2 and ϵ_3 are approximately estimated at 7.09×10^{-4} and 6.66×10^{-4} , with $F_2(\mathbf{0}, \delta, \alpha) \cong 7.09 \times 10^{-2}$ and $F_3(\mathbf{0}, \delta, \alpha) \cong 6.66 \times 10^{-2}$. In addition, $\gamma = 0.1 \min\{\epsilon_2, \epsilon_3\} = 6.66 \times 10^{-5}$.

The changes in key variables including $n_{dc}^{(n)}$, $n_{da} - n_{dc}^{(n)}$, $d_J(X_p^{(n)} - X_p^{(n-1)})$ and $\gamma^{(n)}$ in the repeat process are illustrated in Fig. 5b. The curves of $n_{dc}^{(n)}$ and $n_{da} - n_{dc}^{(n)}$ demonstrate the trade-off between computational efforts used for improving the coverage of the input space and that used for improving the accuracy in regions where up-to-date Pareto-optimal solutions are located. In the initial phase of the repeat process, values of $d_J(X_p^{(n)} - X_p^{(n-1)})$ are relatively large, indicating that the Pareto set drastically moves in the decision variable space. The majority of extra training data is added to improve the coverage of the input space. As the repeat process continues, values of $d_J(X_p^{(n)} - X_p^{(n-1)})$ decrease overall and the Pareto set tends to be fixed. In the last four repeats where the Pareto set remains unchanged, entire computational efforts are used for adding extra training data to improve the accuracy of Pareto-

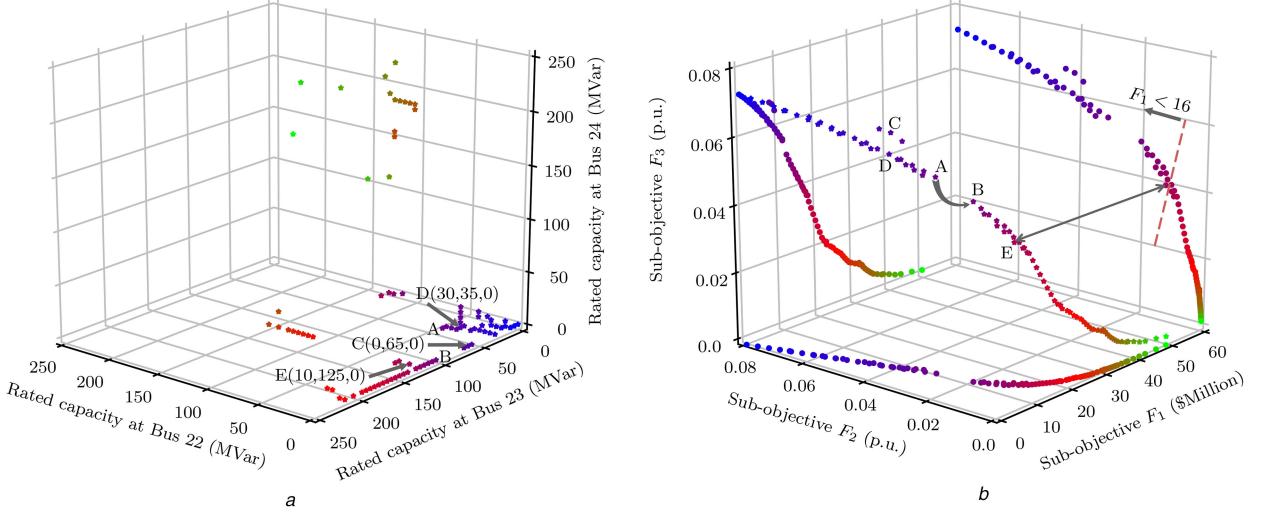


Fig. 6 Solving results of the multi-objective robust dynamic VAR planning model. Each star point corresponds to a Pareto-optimal solution. Also, the Pareto front in (b) is projected onto three coordinate planes, respectively, as shown by round points

(a) Pareto set, (b) Pareto front (star points)

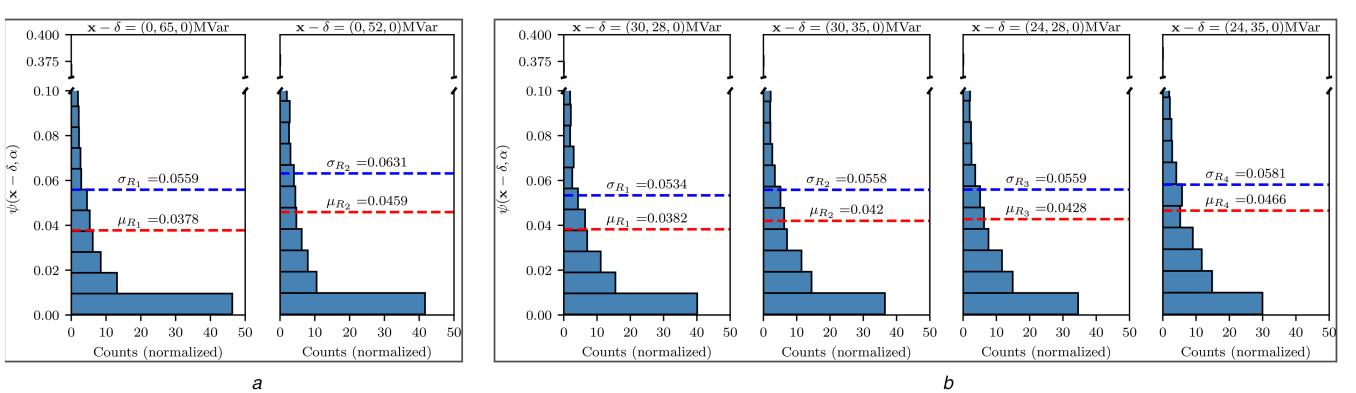


Fig. 7 Histograms of the response ψ of solutions C and D

(a) Histograms of the response ψ of the solution C, (b) Histograms of the response ψ of the solution D

optimal solutions. Finally, the repeat process terminates while $\gamma^{(n)} < \gamma$ with $n = 10$.

Fig. 6 shows solving results of the multi-objective robust dynamic VAR planning model, given by the Pareto set with a total of 87 solutions in Fig. 6a and the corresponding Pareto front in Fig. 6b that is obtained from the last repeat. From Fig. 6b, it can be seen that with the growth of the investment cost, the expectation and robustness of short-term voltage stability level of power grids both improve generally. When the investment cost increases over \$30 million and \$40 million approximately, the improvement in sub-objectives F_2 and F_3 becomes negligible to some extent, respectively. Also, thus, from the point of decision makers, it is inefficient to invest more than \$30 million in the installation of STATCOMs for improving the expectation of the short-term voltage stability level and more than \$40 million for improving the robustness of the short-term voltage stability level.

Additionally, there is an obvious gap between the solutions A and B in the Pareto front in Fig. 6b, where a small increase in the investment cost results in significant improvement in sub-objective F_3 and especially F_2 . This gap is caused by that from solution A to B, the number of uncertainty scenarios, i.e. combinations of δ and α , where severe contingencies lead to short-term voltage instability, dramatically decreases. Thus for decision makers, it is advisable to increase the investment cost to take planning solutions beneath the solution B, provided that the restriction of the investment cost permits.

According to the projection of the Pareto front onto the coordinate plane $F_2 - F_3$ in Fig. 6b, it can be found that on the two sides of the Pareto front, the expectation and robustness of the short-term voltage stability level can be improved simultaneously. However, the improvement of these two objectives is locally

contradictory in the middle part of the Pareto font and further trade-offs between them are needed apart from that between F_1 and F_2 , and F_1 and F_3 . As an illustration of trade-offs between the expectation and robustness of short-term voltage stability level, consider the solutions C and D as shown in Fig. 6. Also, histograms of the response ψ of these two solutions are given in Fig. 7. These two solutions are with the same investment cost equal to \$8.45 million, and values of the sub-objective F_2 of them are $(0.0378 + 0.0459)/2 = 0.04185$ and $(0.0382 + 0.0466)/2 = 0.0424$, respectively. As for the sub-objective F_3 , $F_3 = (0.0559 + 0.0631)/2 + 0.0631 - 0.0559 = 0.0667$ for the solution C and $F_3 = (0.0534 + 0.0581)/2 + 0.0581 - 0.0534 = 0.06045$ for the solution D. Therefore, compared with the solution D, the improvement of the expectation of the short-term voltage stability level of the solution C leads to the weaken robustness of the short-term voltage stability level. Also, the solution C is superior to solution D if the expectation is preferable for decision makers and vice versa.

A merit of considering the robustness of planning solutions w.r.t. the short-term voltage stability level in multi-objective dynamic VAR planning can also be demonstrated by taking the solutions C and D as an example. In the case where only the investment cost and the mean performance of planning solutions are modelled as objectives, the solution D is not a Pareto-optimal solution since the solution C exhibits better mean performance but with the same investment cost compared with D. Also, decision makers are offered no information about the robustness of these two solutions. In fact, however, the disparity of the expectancy measure between these two solutions is fractional but the solution

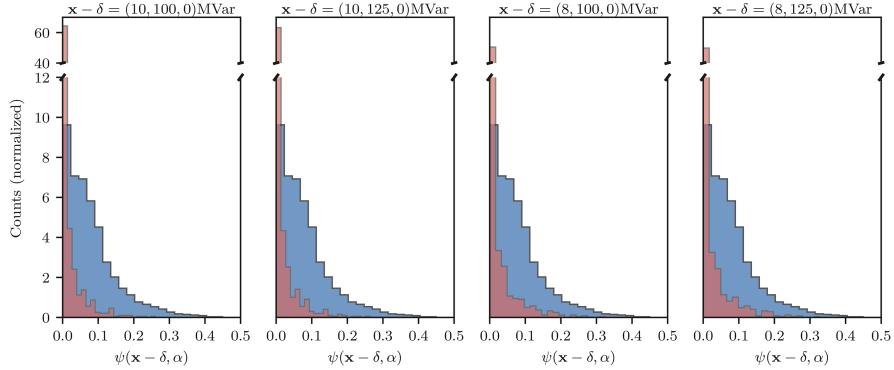


Fig. 8 Histograms of the response ψ of the solution E represented by the red one. Also, the blue histograms represent the case where no STATCOM is installed

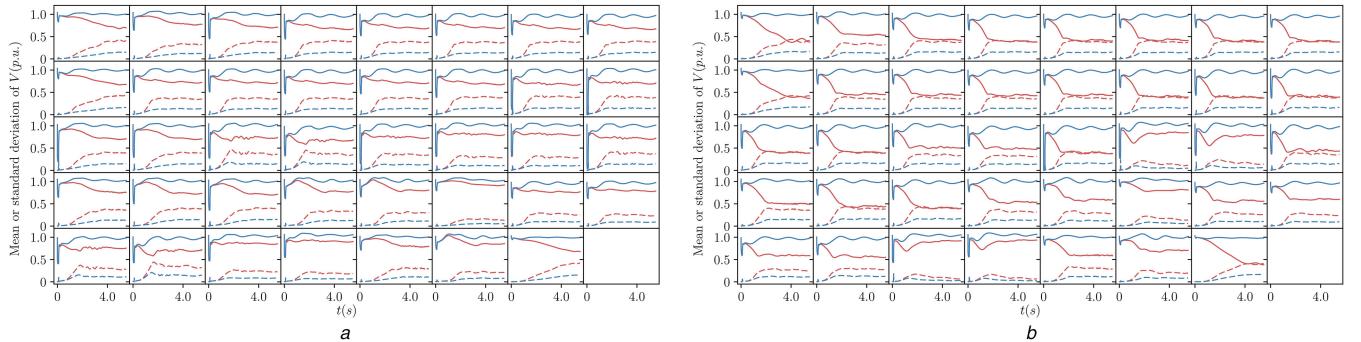


Fig. 9 Statistical time-domain voltage performance of the system with (blue lines) and without (red lines) STATCOMs in the solution E installed. From left to right and top to bottom, sub-figures are statistical $V - t$ graphs of bus 1 to bus 39, respectively. Points on the solid and dashed lines correspond to the mean and standard deviation of voltages of all uncertain scenarios at that time. And uncertain scenarios are the same as that used for uncertainty quantification

(a) Statistical $V - t$ graphs of the contingency occurring in branch 16–17, (b) Statistical $V - t$ graphs of the contingency occurring in branch 21–22

Table 3 Computation time of each main process of the proposed solving methodology

Process	Computation time (hour)
model reduction	0.3226
initial training data sampling	17.5627
ELM structure selection	0.1849
surrogate model construction	0.0264
ϵ -NSGAII optimisation	0.9372
adaptive training data updating	2.1269 (part 1) + 14.6357 (part 2)
total	35.7964

C has an obviously larger dispersion caused by uncertainties α and δ as well, as shown in Fig. 7. Also, if decision makers have no preference between the mean performance and robustness performance of planning solutions, and the investment cost is required to be equal to \$8.45 million, the solution D , instead of the solution C , should be selected as the final one. Thus modelling the robustness performance of planning solutions as an objective in the multi-objective dynamic VAR planning model can avoid optimising the mean performance but leading to planning solutions with an undesirable large dispersion of short-term voltage stability level.

In practice, decision makers generally choose the final solution from the Pareto front according to restrictions of the investment cost, and requirements or preferences of the mean performance and robustness of planning solutions under uncertainties with respect to the short-term voltage stability level. For example, a decision maker claims that investment cost must be strictly <\$16 million, the expectancy measure and dispersion measure should be minimum, and robustness of planning solutions is preferable provided that F_2 is smaller than 0.02, and then the solution E will be the final solution.

The rated capacity of STATCOMs installed of the solution E is shown in Fig. 6a and histograms of the response ψ after STATCOMs being installed are given by Fig. 8. It can be seen that

with STATCOMs' installation, the distribution of the short-term voltage stability level under considered uncertainties shows a significantly stronger central tendency and becomes skewed towards low values of ψ , compared with that without STATCOMs' installation. Therefore the mean performance and robustness of the short-term voltage stability level are both significantly improved. More intuitively, this improvement can also be observed in the time domain from Fig. 9 which compares the statistical $V - t$ graphs of the system with and without STATCOMs in the solution E installed for the two most severe contingencies occurring in branches 16–17 and 21–22, respectively.

5.4 Analysis of computation time

The computation time of each main process of our proposed solving methodology applied to the IEEE 39-bus system is given in Table 3. Processes with fractional computation time, such as computing of stop condition, are omitted here. In Table 3, the process of adaptive training data updating is divided into two parts: part 1 including line 1 to line 15 in Algorithm 1, and part 2 including line 16 to the end of Algorithm 1.

Assume that the planning model is solved by evaluating ψ in the optimisation process directly, instead of the ELM model $\tilde{\psi}$. A total of 10^4 evaluations of objective functions are assumed to be needed, which are actually not enough for multi-objective algorithms to obtain the satisfied Pareto-optimal solutions. Also, $n_s = 10^2$ that is also not large enough, is used for uncertainty quantification. The uncertainties δ and computation time of multi-objective optimisation excluding the evaluation of F_2 and F_3 are both omitted. With that computation time of one evaluation of ψ is 8.4301 s, the total computation time is $8.4301 \text{ s} \times 10^2 \times 10^4 = 2342.017 \text{ h}$ in this case. Thus, the proposed solving methodology with the application of surrogate modelling technologies significantly reduced the computing burden that is unaffordable with the personal computer used in this work and makes the solving process computationally efficient.

Another important issue of our proposed solving methodology is its computation time for large-scale power grids. In Table 3, the computation time of initial training data sampling and part 2 of adaptive training data updating together accounts for the majority of total computation time. Approximately, there is a linear relationship between computation time of the entire process and that of one evaluation of ψ . Also, for a real 220 kV power grid with 205 buses and 31,926 nodes being processed in time-domain simulations, which is reported in our previous work [15], the computation time of one evaluation of ψ is 98.6424 s. The total computation time of this large-scale power grid is roughly $35.7964 \times 98.6424 / 8.4301 = 418.8613$ h, and which can easily become acceptable by using some powerful computing technologies, such as parallel computing and distributed computing when performing time-domain simulations.

6 Conclusion

In this study, a novel multi-objective robust dynamic VAR planning approach is proposed. The dynamic VAR planning problem is formulated as a multi-objective optimisation model with objectives including the investment cost, the expectation, and robustness of the short-term voltage stability level. Uncertainties in the peak load level, the maximum proportion of dynamic load, fault clearing time and deviation of the actual capacity of dynamic VAR compensators from rated capacity when contingencies occur are considered in this formulation. The complexity of the planning model is first reduced by selecting severe contingencies and potential buses. Latin hypercube sampling is used for the uncertainty quantification. Also, then the simplified multi-objective optimisation model is solved by the combination of ϵ -NSGAII and ELM-based surrogate modelling with adaptive training data sampling.

Numerical experiments verify the proposed approach. The application of surrogate modelling technologies significantly reduces the unaffordable computing burden and makes the solving process computationally efficient. By solving the planning model, decision makers can be offered the Pareto set and Pareto front for further analysis and make trade-offs among the investment cost, the expectation and robustness of the short-term voltage stability level. In addition, by modelling the robustness performance of planning solutions as an objective, it can be avoided to optimise the mean performance but leading to planning solutions with an undesirable large dispersion of the short-term voltage stability level.

The present approach only considers the improvement of the short-term voltage stability level by installing dynamic VAR compensators which, however, can also offer fast-acting damping to efficiently improve transient stability level. Thus dynamic VAR planning considering both improvement of short-term voltage stability level and transient stability level under uncertainties is a future work.

7 Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (51777067 and 51407069), in part by the Fundamental Research Funds for the Central Universities (2016YQ02), and in part by the Faculty of Engineering and IT Mid-Career Researcher Development Scheme, University of Sydney.

8 References

- [1] Andersson, G., Donalek, P., Farmer, R., et al.: 'Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance', *IEEE Trans. Power Syst.*, 2005, **20**, (4), pp. 1922–1928
- [2] Romero, J.J.: 'Blackouts illuminate India's power problems', *IEEE Spectr.*, 2012, **49**, (10), pp. 11–12
- [3] Kundur, P., Paserba, J., Ajjarapu, V., et al.: 'Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions', *IEEE Trans. Power Syst.*, 2004, **19**, pp. 1387–1401
- [4] de Leon, J.A.D., Taylor, C.W.: 'Understanding and solving short-term voltage stability problems'. 2002 IEEE Power Engineering Society Summer Meeting, July 2002, vol. 2, pp. 745–752
- [5] Noroozian, M., Petersson, N.A., Thorvaldson, B., et al.: 'Benefits of SVC and STATCOM for electric utility application'. Transmission and Distribution Conf. and Exposition, 2003 IEEE PES, September 2003, vol. 3, pp. 1143–1150
- [6] Hingorani, N.G., Gyugyi, L.: 'Static shunt compensators: SVC and STATCOM' (Wiley-IEEE Press, Hoboken, NJ, USA, 2000), pp. 135–207
- [7] Sapkota, B., Vittal, V.: 'Dynamic var planning in a large power system using trajectory sensitivities', *IEEE Trans. Power Syst.*, 2010, **25**, pp. 461–469
- [8] Tiwari, A., Ajjarapu, V.: 'Optimal allocation of dynamic VAR support using mixed integer dynamic optimization', *IEEE Trans. Power Syst.*, 2011, **26**, pp. 305–314
- [9] Paramasivam, M., Salloum, A., Ajjarapu, V., et al.: 'Dynamic optimization based reactive power planning to mitigate slow voltage recovery and short term voltage instability', *IEEE Trans. Power Syst.*, 2013, **28**, pp. 3865–3873
- [10] Huang, W., Sun, K., Qi, J., et al.: 'Voronoi diagram based optimization of dynamic reactive power sources'. 2015 IEEE Power Energy Society General Meeting, July 2015, pp. 1–5
- [11] Xu, Y., Dong, Z.Y., Meng, K., et al.: 'Multi-objective dynamic VAR planning against short-term voltage instability using a decomposition-based evolutionary algorithm', *IEEE Trans. Power Syst.*, 2014, **29**, pp. 2813–2822
- [12] Huang, W., Sun, K., Qi, J., et al.: 'Optimal allocation of dynamic VAR sources using the Voronoi diagram method integrating linear programming', *IEEE Trans. Power Syst.*, 2017, **32**, (6), pp. 4644–4655
- [13] Wildenhues, S., Rueda, J.L., Erlich, I.: 'Optimal allocation and sizing of dynamic VAR sources using heuristic optimization', *IEEE Trans. Power Syst.*, 2015, **30**, pp. 2538–2546
- [14] Qi, J., Huang, W., Sun, K., et al.: 'Optimal placement of dynamic VAR sources by using empirical controllability covariance', *IEEE Trans. Power Syst.*, 2017, **32**, (1), pp. 240–249
- [15] Han, T., Chen, Y., Ma, J., et al.: 'Surrogate modeling based multi-objective dynamic VAR planning considering short-term voltage stability and transient stability', *IEEE Trans. Power Syst.*, 2018, **33**, (1), pp. 622–633
- [16] Xu, Y., Dong, Z.Y., Xiao, C., et al.: 'Optimal placement of static compensators for multi-objective voltage stability enhancement of power systems', *IET Gener. Transm. Distrib.*, 2015, **9**, (15), pp. 2144–2151
- [17] Tahboub, A.M., ElMoursi, M.S., Woon, W.L., et al.: 'Multi-objective dynamic VAR planning strategy with different shunt compensation technologies', *IEEE Trans. Power Syst.*, 2017, **PP**, (99), pp. 1–1, DOI: 10.1109/TPWRS.2017.2751080
- [18] Douglas, A.P., Breipohl, A.M., Lee, F.N., et al.: 'Risk due to load forecast uncertainty in short term power system planning', *IEEE Trans. Power Syst.*, 1998, **13**, (4), pp. 1493–1499
- [19] Billinton, R., Huang, D.: 'Effects of load forecast uncertainty on bulk electric system reliability evaluation', *IEEE Trans. Power Syst.*, 2008, **23**, (2), pp. 418–425
- [20] De la Torre, T., Feltes, J.W., San Roman, T.G., et al.: 'Deregulation, privatization, and competition: transmission planning under uncertainty', *IEEE Trans. Power Syst.*, 1999, **14**, (2), pp. 460–465
- [21] Vaahedi, E., Li, W., Chia, T., et al.: 'Large scale probabilistic transient stability assessment using BC Hydro's on-line tool', *IEEE Trans. Power Syst.*, 2000, **15**, (2), pp. 661–667
- [22] Silva, I.D.J., Rider, M.J., Romero, R., et al.: 'Transmission network expansion planning considering uncertainty in demand', *IEEE Trans. Power Syst.*, 2006, **21**, (4), pp. 1565–1573
- [23] El Moursi, M.S., Zeineldin, H.H., Kirtley, J.L., et al.: 'A dynamic master/slave reactive power-management scheme for smart grids with distributed generation', *IEEE Trans. Power Deliv.*, 2014, **29**, (3), pp. 1157–1167
- [24] Wilamowski, B.M., Irwin, J.D.: 'Power electronics and motor drives' (CRC Press, Boca Raton, FL, USA, 2016)
- [25] Juliano, E., Pérez, E.A.: 'Application of surrogate-based global optimization to aerodynamic design' (Springer, New York, NY, USA, 2015)
- [26] Kollat, J.B., Reed, P.M.: 'Comparing state-of-the-art evolutionary multiobjective algorithms for long-term groundwater monitoring design', *Adv. Water Resour.*, 2006, **29**, (6), pp. 792–807
- [27] Reed, P.M., Minsker, B.S., Goldberg, D.E.: 'Simplifying multiobjective optimization: an automated design methodology for the nondominated sorted genetic algorithm', *Water Resour. Res.*, 2003, **39**, (7), pp. 1–5
- [28] Huang, G., Zhu, Q., Siew, C.: 'Extreme learning machine: theory and applications', *Neurocomputing*, 2006, **70**, (1), pp. 489–501
- [29] Akusok, A., Björk, K.-M., Miche, Y., et al.: 'High-performance extreme learning machines: a complete toolbox for big data applications', *IEEE Access*, 2015, **3**, pp. 1011–1025
- [30] Levandowsky, M., Winter, D.: 'Distance between sets', *Nature*, 1971, **234**, (5323), pp. 34–35
- [31] Pai, A.: 'Energy function analysis for power system stability' (Springer Science & Business Media, New York, NY, USA, 2012)
- [32] Hadka, D.: 'MOEA framework user guide. Version 2.7', 2015
- [33] Kundur, P.: 'Power system stability and control' (McGraw-Hill, New York NY, USA, 1994)
- [34] Ying, Y., Tang, Y., Li, B.: 'User manual of PSD-BPA stability programs' (China Electric Power Research Institute, Beijing, China, 2012)
- [35] Matala, A.: 'Sample size requirement for Monte Carlo simulations using Latin hypercube sampling' (Helsinki University of Technology, Department of Engineering Physics and Mathematics, Systems Analysis Laboratory, Espoo, Finland, 2008)