Analytical Representation of Data-driven Transient Stability Constraint and Its Application in Preventive Control

Yiwei Fu, Xiaohua Zhang, Lei Chen, Zengyao Tian, Kaiyuan Hou, and Haikuan Wang

***Abstract*—Accurate transient stability assessment (TSA) and effective preventive control are important for the stable operation of power systems. With the superiorities in precision and efficiency, data-driven methods are widely used in TSA nowadays. Data-driven TSA model can be adopted in the stability constraints of preventive control optimization, but existing methods are mostly iteration-based ones, which may result in low efficiency, sometimes even non-convergence. In this paper, an analytical representation method of data-driven transient stability constraint is proposed based on a non-parametric regression model built for TSA. Key feature extraction and dominant sample selection are proposed to reduce the scale of the TSA model, and bi-level linearization is applied to further modify it. Optimal preventive control model is then formulated as a mixed-integer linear program (MILP) problem with the linearized analytical data-driven transient stability constraint, which can be solved without iterations. An overall procedure of data-driven TSA and preventive control is finally developed. Case studies show that the proposed method has high accuracy in TSA and can achieve effective preventive control of power system with high efficiency.**

***Index Terms*—Data-driven, critical clearing time, transient stability constraint, non-parametric regression, optimal preventive control.**

I. Introduction

TRANSIENT stability assessment (TSA) is a critical means to monitor the operation of modern power system [1], which can determine the stability and stability margin of the pre-fault system in case of a specific contingency [2]. Many transient stability indices (TSIs) can be used to quantify the transient stability margin, and the critical clearing time (CCT) is one of the most commonly used TSIs [3]. When the TSI of a power system obtained by TSA exceeds the preset stability range, the system is identified as unstable. At this time, preventive control is necessary for power system to maintain stable operation [4]-[6]. Preventive control generally refers to adjusting the steady-state operation point of the pre-fault power system through efficient calculation, so that the system can withstand the occurrence of the specific contingency [7].

The preventive control of transient stability is usually realized by solving the transient stability constrained-optimal power flow (TSC-OPF) [8]. By adding the stability constraints to the optimal power flow model, preventive control can achieve the coordination between economics and stability requirements of power system [9]. However, the high nonlinearity of power system transient stability makes the calculation of TSC-OPF difficult in two aspects [10]: ① how to construct the transient stability constraints that need to be added to the conventional OPF model; ② how to solve the optimization problem with these transient stability constraints.

The conventional methods for TSC-OPF can be classified as dynamic optimization methods, simplification-based methods, and meta-heuristics methods according to [7], [11].

Simultaneous discretization [12], constraint transcription [13], and multiple shooting [14] are three widely-used dynamic optimization methods for solving TSC-OPF problems. As for simultaneous discretization, [15] and [16] discretize all the differential equations in the transient stability constraints to non-linear algebraic equations. With gradient calculation using trajectory sensitivities [17], constraint transcription methods can decouple the optimization algorithms and simulation tools [18]. Multiple shooting is introduced to solve TSC-OPF problems as it is a hybridization of simultaneous discretization and constraint transcription [19]. However, once the system scale increases, these three kinds of dynamic optimization methods will face high computational burden and thus lose their efficiency and practicability.

The simplification-based methods introduce the direct TSA methods such as single machine equivalent to the TSC-OPF problems [20]. These methods may face convergence issues during the TSC-OPF iterations and suffer from inconsistent solution in practical use [11].

Manuscript received: August 14, 2020; revised: November 15, 2020; accepted: July 8, 2021. Date of CrossCheck: July 8, 2021. Date of online publication: July 28, 2021.

This article was supported by National Key R&D Program of China (No. 2018YFB0904500) and State Grid Corporation of China (No. SGLNDK00KJJS1800236).

Y. Fu and L. Chen (corresponding author) are with State Key Laboratory of Control and Simulation of Power System and Generation Equipment, Department of Electrical Engineering, Tsinghua University, Beijing 100084, China (e-mail: fuyw19@mails.tsinghua.edu.cn; chenlei08@tsinghua.edu.cn).

X. Zhang is with State Grid Jibei Electric Power Company, Beijing 100053, China (e-mail: zhang-xiaohua@sgcc.com.cn).

Z. Tian, K. Hou, and H. Wang are with Northeast China Branch of State Grid Corporation of China, Shenyang 110180, China (e-mail: houkaiyuan2010@163.com; 523651872@qq.com; 20545087@qq.com.)

DOI: 10.35833/MPCE.2020.000608

The meta-heuristics methods are problem-independent, and their applications for solving TSC-OPF include the particle swarm optimization (PSO) algorithm [21], non-dominated sorting genetic algorithm II (NSGA-II) [22], genetic algorithm [23] and so on. These algorithms often require a considerable number of iterations and cannot guarantee the optimality and applicability of results.

Comparing with the aforementioned conventional preventive transient stability control methods, data-driven methods have clear superiority in accuracy and efficiency, and thus have been applied in power system stability in recent years [24]. A basic idea of using data-driven methods for preventive control is to build a data-driven TSA model first, then construct the transient stability constraints based on it, and solve the TSC-OPF model including the data-driven transient stability constraints.

There are many data-driven methods such as artificial neural network (ANN) [25], support vector machine (SVM) [26], decision tree [27], *K*-nearest neighbor [28], and deep learning [29], [30] have been applied to implement TSA in the power system. Lots of research has shown that data-driven methods can achieve fast and accurate TSA, which leads to its rapid development.

When the TSA result shows that the power system is transient unstable, the preventive control is essential. However, there are few researches on how to implement preventive control if the stability constraint is composed of a data-driven TSA model and a preset stability threshold [7]. The existing preventive control methods proposed in relevant research are mostly iteration-based, which can be classified into four schemes.

1) The iterative sensitivity method [31]-[34]. The sensitivities of the TSI to the control variables (usually the outputs of generators) are calculated by the information provided by the data-driven or linearized TSA model. Then, according to the value of the corresponding sensitivity sorted from large to small, the control variables are iteratively adjusted until the system is stable.

2) The iterative gradient method [35]. The gradient descent method is one of the most fundamental iterative optimization methods for unconstrained optimization problem. When it is used in preventive control, the stability constraints directly formed by the data-driven TSA model and other constraints all need to be put into the loss function.

3) The iterative probability method. Bayesian optimization, which uses the pattern with the highest probability in each iteration, is adopted to solve the preventive control problem in [36]. Probability is also estimated in each iteration in [37]. Reference [38] finds the nearest secure pattern in each iteration to test if power flow is converged, which also can be classified into iterative probability method.

4) The intelligent optimization method [39]. Stochastic search algorithms based on biological intelligence or physical phenomena, which are called intelligent optimization methods, have also been used to solve preventive control strategy.

However, these data-driven control methods have limitations in four aspects: ① the convergence of the optimization algorithms in these methods cannot be guaranteed, and the obtained solution may not be optimal [40]; ② when there are too many constraints in the optimal control model, it is difficult to design the iterative algorithm flowchart; ③ most of the intelligent optimization algorithms in these methods lack the support of mathematical theory; ④ the searching rates of these methods are usually slow, which leads to low efficiency [41].

The source of these limitations mainly lies in the lack of transparency in the existing data-driven models, which makes it difficult to summarize the analytical expression of data-driven stability constraints. Without analytical expression, the stability-constrained preventive control model cannot be directly solved by frequently-used optimization solvers, and the mechanism of stability constraints is hard to analyze. If the analytical derivation of the data-driven stability constraints is realized, it will not only lay the foundation for the unified solution of the preventive control optimization model, but also provide the possibility to analyze the mechanism of the preventive control model. Therefore, the analytical representation method of data-driven transient stability constraint and its application in preventive control are proposed in this paper.

The premise of realizing the analytical representation of stability constraints is to choose the appropriate data-driven TSA method. The data-driven methods generally fall into two categories [42]: parametric regression and non-parametric regression. The general form of non-parametric model is a regression expression of the labels of all training samples. Compared with parametric model, non-parametric model is easier to analyze the influence of each training sample on the final result [43]. Therefore, this paper adopts non-parametric regression to construct the TSA model, and then modifies it to form the analytical representation of data-driven transient stability constraints according to the requirements of preventive control, so as to realize the efficient solution of the preventive control optimization model.

The main contributions of this paper are as follows.

1) The analytical representation of data-driven transient stability constraint is first proposed based on a non-parametric regression based TSA model. Key feature extraction and dominant sample selection are proposed to reduce the scale of the analytical stability constraints. The data-driven transient stability constraint is linearized as a mixed-integer linear program (MILP) problem.

2) An optimal preventive control (OPC) model based on the proposed analytical stability constraints is innovatively built and efficiently solved. The hot-start linear power flow equations [44] are adopted as the power flow equality constraint.

3) A novel procedure of data-driven TSA and OPC is presented. The solution of the OPC problem can be obtained by solving a programming problem without iterations.

The rest of the paper is organized as follows. Section II introduces the formulation of data-driven OPC and the necessity of analytic processing. In Section III, the analytical representation of transient stability constraint is proposed, in which the construction and modifications of TSA model are

elaborated. In Section IV, the procedure of data-driven TSA and OPC is illustrated. In Section V, case studies are presented to validate the effectiveness of the proposed method. The conclusion and the future work are drawn in Section VI.

II. Formulation of Data-driven OPC

A. Data-driven TSA

The construction of data-driven TSA model is the foundation of data-driven preventive control. For a power system with fixed network topology, its transient stability can be determined by the contingency and the pre-fault power flow, and therefore the mapping can be expressed as:

$$\begin{cases} S = f_{\Delta}(\boldsymbol{p}) \\ \Delta \in \boldsymbol{B} \end{cases} \tag{1}$$

where S is the selected TSI; $\boldsymbol{p}$ is the pre-fault power flow; and Δ represents the contingency, which is generally considered from a specific set $\boldsymbol{B}$. Equation (1) also indicates that the mapping $f_{\Delta}(\cdot)$ between $\boldsymbol{p}$ and S is different under different contingencies.

With massive simulation data accumulated from practical TSA, various machine-learning algorithms can be applied to train the mapping (1), and the data-driven TSA model is thus obtained.

B. OPC Model

For transient stability, the mathematical nature of preventive control is the TSC-OPF problem [7]. Combined with the data-driven TSA model, the TSC-OPF model can be described in a concise form by:

$$\min C(\boldsymbol{x}, \boldsymbol{u}) \tag{2}$$

s.t.

$$g_s(\boldsymbol{x}, \boldsymbol{u}) = 0 \tag{3}$$

$$h_s^- \le h_s(\boldsymbol{x}, \boldsymbol{u}) \le h_s^+ \tag{4}$$

$$S = f_{\Delta}(\boldsymbol{p}) \in K \tag{5}$$

where $\boldsymbol{u}$ is a set of control variables, including the active power and voltage magnitude square of the generator bus, and the voltage angle and voltage magnitude square of the swing bus; $\boldsymbol{x}$ is a set of dependent variables, including the active and reactive power of the swing bus, the voltage angle and reactive power of the generator bus, and the voltage angle and voltage magnitude square of the load bus; $C(\boldsymbol{x}, \boldsymbol{u})$ is the objective function; $g_s(\boldsymbol{x}, \boldsymbol{u})$ is the set of equality constraints, which usually contains the nodal power balance equations and the power flow equations in a specific operation scenario; $h_s(\boldsymbol{x}, \boldsymbol{u})$ is the set of inequality constraints, which represents the limits of the control variables and the power system operation state; and K is the preset stability requirement. A simplified expression of the data-driven transient stability constraints is shown as (5), in which $S = f_{\Delta}(\boldsymbol{p})$ is the explicit data-driven expression of TSI.

1) Objective Function

The goal of preventive control is to minimize the regulation of generator while adjusting the power system to transient stability. As the transient stability of the system is guaranteed in the form of constraints, the objective function of the control model is to minimize the regulation of generator output relative to the initial operation mode.

$$\min C(\boldsymbol{x}, \boldsymbol{u}) = \sum_{g} (P_{G,g} - P_{G,g}^0)^2 \tag{6}$$

where $P_{G,g}$ is the active power output of generator g in the current operation scenario; and $P_{G,g}^0$ is the output of generator g in the original operation scenario.

2) Steady-state Equality Constraints

The steady-state equality constraints (3) consist of nodal power balance constraint and power flow equality constraint. Considering the reactive power injection of grounded capacitors, the nodal power balance constraint is expressed as:

$$\sum_{g \in G_i} P_{G,g} - P_{L,i} = \sum_{(i,j) \in K_i} P_{ij} + V_i^2 \sum_{j=1}^{N} g_{ij} \quad i \in S_N \tag{7}$$

$$\sum_{g \in G_i} Q_{G,g} - Q_{L,i} = \sum_{(i,j) \in K_i} Q_{ij} + V_i^2 \sum_{j=1}^{N} (-b_{ij}) + Q_i^{sc} \quad i \in S_N \tag{8}$$

where S_N is the collection of total buses; $P_{G,g}$ and $Q_{G,g}$ are the active and reactive outputs of the generator g, respectively; $P_{L,i}$ and $Q_{L,i}$ are the active and reactive loads of bus i, respectively; P_{ij} and Q_{ij} are the active and reactive power flows of line (i,j), respectively; g_{ij} and b_{ij} are the conductance and susceptance of line (i,j), respectively; Q_i^{sc} is the reactive power injection of grounded capacitor i; V_i is the voltage amplitude of bus i; G_i is the set of generators connected to bus i; and K_i is the set of lines connected to bus i.

The power flow equality constraint is generally formulated as:

$$P_{ij} = g_{ij}(V_i^2 - V_i V_j \cos\theta_{ij}) - b_{ij} V_i V_j \sin\theta_{ij} \tag{9}$$

$$Q_{ij} = -b_{ij}(V_i^2 - V_i V_j \cos\theta_{ij}) - g_{ij} V_i V_j \sin\theta_{ij} \tag{10}$$

where V_j is the voltage amplitude of bus j; and θ_{ij} is the phase angle difference between bus i and bus j. Since this constraint is non-linear, the method in [44] is adopted to linearize it so as to obtain an analytical expression which can help to realize the efficient solution of the OPC model. The specific linearized expression will be introduced in Section IV-A.

3) Steady-state Inequality Constraints

The upper and lower limits of voltage square, active power, and reactive power of generators typically constitute the steady-state inequality constraints (4) as given by:

$$\left(V_i^{\min}\right)^2 \le V_i^2 \le \left(V_i^{\max}\right)^2 \quad i \in S_N \tag{11}$$

$$P_{g,i}^{\min} \le P_{g,i} \le P_{g,i}^{\max} \quad i \in S_G \tag{12}$$

$$Q_{g,i}^{\min} \le Q_{g,i} \le Q_{g,i}^{\max} \quad i \in S_G \tag{13}$$

where S_G is the collection of generator buses; superscripts max and min represent the maximum and minimum values of the corresponding variable, respectively; and $P_{g,i}$ and $Q_{g,i}$ are the active and reactive power outputs of generator i, respectively.

4) Data-driven Transient Stability Constraint

The traditional transient stability constraint in power system is expressed by differential-algebraic equations, which

are often solved by time-domain simulation method, only at the cost of long calculation time.

The data-driven transient stability constraints proposed in this paper can be expressed as a whole in (5) and further divided into the following two formulas:

$$S=f_{\Delta}(\boldsymbol{p}) \tag{14}$$

$$S^{*}=f_{\Delta}^{*}(\boldsymbol{p}) \in K \tag{15}$$

These two formulas show that the construction of transient stability constraints needs to go through two steps: ① the expression (14) of the selected TSI S is obtained by a non-parametric regression method; ② the trained TSA model (14) is modified to obtain the approximate expression $S^{*}=f_{\Delta}^{*}(\boldsymbol{p})$ of TSI in (15) according to the requirement of preventive control. The stability range K is set for the final stability index S^{*}. The expression (15) obtained finally is the analytical representation of transient stability constraint therefrom.

C. Significance of Analytical Representation

Instead of the traditional transient stability constraints, the analytical transient stability constraints can avoid the repeated calculation of time-domain simulation in the process of solving the OPC problem, thus greatly saving calculation time. In addition, the estimation accuracy of data-driven TSA model can be guaranteed with sufficient training data. Therefore, as long as the method of modifying the TSA model to the stability constraints is reasonable and effective, the analytical transient stability constraints can also maintain high accuracy.

The acquisition of a concise and accurate analytical stability constraint enables the general optimizer to solve the preventive control problem directly, which greatly improves the efficiency of optimization. Meanwhile, since the transient stability constraint ultimately exists in a definite expression, other constraints do not need to be satisfied by iteration.

III. Analytical Representation of Transient Stability Constraint

There are two key points to obtain the analytical expression of transient stability constraint: ① select an appropriate data-driven method to build the TSA model as shown in (14); and ② make reasonable and effective modifications to obtain a refined and accurate analytical expression of TSI as in (15).

A. Data-driven TSA Model Based on Kernel Regression

Since the non-parametric regressions do not presuppose the mathematic form of function, they have better performance than the parametric methods in complex problems such as TSA. The kernel regression, one of the most popular non-parametric methods, can identify the similarity between the labeled data and the data to be estimated by the inner product in the kernel function, and then determine the contribution of each training sample in the final estimation result. The obtained kernel regression has clear physical significance and explicit expression, which lays a foundation for its analytical representation in preventive control.

In this paper, CCT is taken as the TSI, and the MD-kernel regression method proposed in [42], [45] which combines the Nadaraya-Watson kernel regression and the Mahalanobis distance, is adopted to construct the TSA model. The target of this method is to determine the mapping as:

$$\hat{y}_{\mathrm{CCT}}(\boldsymbol{X})=f_{\Delta}(\boldsymbol{X}) \tag{16}$$

where $\hat{y}_{\mathrm{CCT}}(\boldsymbol{X})$ is the estimated CCT value with the pre-fault power flow $\boldsymbol{X}$.

Since the Mahalanobis distance can consider the correlation among different dimensions of variables, the MD-kernel regression model proposed in [42] is constructed by adopting the Mahalanobis distance to replace the Euclidean distance in the conventional Nadaraya-Watson kernel regression function.

The specific expression of the Mahalanobis distance is:

$$D_{\mathrm{M}}(\boldsymbol{x}_i,\boldsymbol{x})=\sqrt{(\boldsymbol{x}_i-\boldsymbol{x})^{\mathrm{T}}\boldsymbol{M}(\boldsymbol{x}_i-\boldsymbol{x})} \tag{17}$$

where $\boldsymbol{M}$ is the Mahalanobis distance metric matrix; and $D_{\mathrm{M}}(\boldsymbol{x}_i,\boldsymbol{x})$ is the dissimilarity between two input scenarios $\boldsymbol{x}_i$ and $\boldsymbol{x}$ when applied in TSA [42].

The MD-kernel regression constructed by $D_{\mathrm{M}}(\boldsymbol{x}_i,\boldsymbol{x})$ can be expressed as:

$$\kappa_{\mathrm{MD}}(\boldsymbol{x}_i,\boldsymbol{x})=\exp\left(-\frac{\gamma}{2}D_{\mathrm{M}}^{2}(\boldsymbol{x}_i,\boldsymbol{x})\right) \tag{18}$$

where γ is a smoothing parameter.

By substituting (18) into the general Nadaraya-Watson kernel regression function, the final expression of transient stability estimator based on MD-kernel regression is obtained as:

$$\hat{y}_{\mathrm{CCT}}(\boldsymbol{X})=\frac{\sum_{i=1}^{N^{\mathrm{train}}}\kappa_{\mathrm{MD}}(\boldsymbol{X}_i^{\mathrm{train}},\boldsymbol{X})Y_i^{\mathrm{train}}}{\sum_{i=1}^{N^{\mathrm{train}}}\kappa_{\mathrm{MD}}(\boldsymbol{X}_i^{\mathrm{train}},\boldsymbol{X})}=\sum_{i=1}^{N^{\mathrm{train}}}\mu_i Y_i^{\mathrm{train}} \tag{19}$$

$$\mu_i=\frac{\exp\left(-\frac{\gamma}{2}D_{\mathrm{M}}^{2}(\boldsymbol{X}_i^{\mathrm{train}},\boldsymbol{X})\right)}{\sum_{i=1}^{N^{\mathrm{train}}}\exp\left(-\frac{\gamma}{2}D_{\mathrm{M}}^{2}(\boldsymbol{X}_i^{\mathrm{train}},\boldsymbol{X})\right)} \tag{20}$$

where N^{train} is the total number of training samples; $\boldsymbol{X}_i^{\mathrm{train}}$ is the input vector of the i^{th} training sample; Y_i^{train} is the CCT of the i^{th} training sample; $\boldsymbol{X}$ is the input vector of the sample to be estimated; and μ_i is the weighting coefficient of the CCT value of the i^{th} training sample on estimated result $\hat{y}_{\mathrm{CCT}}(\boldsymbol{X})$.

The core issue of obtaining the final expression (19) is to determine the distance $D_{\mathrm{M}}(\boldsymbol{X}_i^{\mathrm{train}},\boldsymbol{X})$, that is to find the optimal $\boldsymbol{M}$ in (17). Inspired by [42], the objective function of the optimization model to obtain $\boldsymbol{M}$ can be summarized as:

$$\min l(\boldsymbol{M})=g(\boldsymbol{M})+h(\boldsymbol{M}) \tag{21}$$

$$g(\boldsymbol{M})=\sum_{i=1}^{N^{\mathrm{train}}}\left[\frac{\sum_{j=1,j\neq i}^{N^{\mathrm{train}}}\kappa_{\mathrm{MD}}(\boldsymbol{X}_i^{\mathrm{train}},\boldsymbol{X}_j^{\mathrm{train}})y_{\mathrm{CCT},j}^{\mathrm{train}}}{\sum_{j=1,j\neq i}^{N^{\mathrm{train}}}\kappa_{\mathrm{MD}}(\boldsymbol{X}_i^{\mathrm{train}},\boldsymbol{X}_j^{\mathrm{train}})}-y_{\mathrm{CCT},i}^{\mathrm{train}}\right]^{2} \tag{22}$$

$$h(\boldsymbol{M})=\alpha\|\boldsymbol{M}\|^{2} \tag{23}$$

where $g(\boldsymbol{M})$ is the leave-one-out error of stability estimation; $h(\boldsymbol{M})$ is the regularization terms of the elements in $\boldsymbol{M}$; $y_{\mathrm{CCT},i}^{\mathrm{train}}$ and $y_{\mathrm{CCT},j}^{\mathrm{train}}$ are the CCT values of the i^{th} and j^{th} training samples, respectively; and α is the amplification coefficient, which is set to be 10^{-8}. The optimization problem shown in (21)-(23) can be solved by the leave-one-out method and the gradient descent algorithm [42].

B. Complexity of Kernel Regression Model

After the data-driven CCT estimator is trained, the threshold of CCT value can be set to define the transient stability constraint, which is added to the preventive control model to make the transient stability of power system meet certain stability margin. Therefore, the transient stability constraint can be preliminarily expressed as:

$$\hat{y}_{\mathrm{CCT}}(\boldsymbol{X}) = \frac{\sum_{i=1}^{N^{\mathrm{train}}} \exp\left(-\frac{\gamma}{2} D_{\mathrm{M}}^2\left(\boldsymbol{X}_i^{\mathrm{train}}, \boldsymbol{X}\right)\right) Y_i^{\mathrm{train}}}{\sum_{i=1}^{N^{\mathrm{train}}} \exp\left(-\frac{\gamma}{2} D_{\mathrm{M}}^2\left(\boldsymbol{X}_i^{\mathrm{train}}, \boldsymbol{X}\right)\right)} \geq y_{\mathrm{thre}} \quad (24)$$

where y_{thre} is a preset stability threshold.

However, the TSA model constructed by data-driven method is non-linear and has certain complexity. If this model is directly inserted into the OPC model as a constraint, it is difficult to achieve the integrated solution. Meanwhile, different from the input variables of the TSA model, the variables are divided into control variables and dependent variables in preventive control. Therefore, only after a series of modifications to the trained TSA model in accordance with the requirements of preventive control, the analytical transient stability constraint can be finally obtained and added into the preventive model.

The complexity and non-linearity of the TSA model mainly come from the exp(·) function and the square of Mahalanobis distance $D_{\mathrm{M}}^2\left(\boldsymbol{X}_i^{\mathrm{train}}, \boldsymbol{X}\right)$ in the MD-kernel regression given by (18), in which the concrete expression of $D_{\mathrm{M}}^2\left(\boldsymbol{X}_i^{\mathrm{train}}, \boldsymbol{X}\right)$ can be obtained by expanding and deducing (17):

$$D_{\mathrm{M}}^2\left(\boldsymbol{X}_i^{\mathrm{train}}, \boldsymbol{X}\right) = \sum_{p=1}^{k} m^{pp}\left(x^p\right)^2 + 2\sum_{p\neq q} m^{pq} x^p x^q - \sum_{p=1}^{k} C_{a,i}^p x^p + C_{b,i} \quad (25)$$

$$C_{a,i}^p = \sum_{v=1}^{k}\left(m^{pv} + m^{vp}\right) x_{\mathrm{train},i}^v \quad (26)$$

$$C_{b,i}^p = \sum_{p=1}^{k}\left(\sum_{v=1}^{k} m^{vp} x_{\mathrm{train},i}^v\right) x_{\mathrm{train},i}^p \quad (27)$$

where m^{pq} is the element at row p and column q in the matrix $\boldsymbol{M}$; m^{pv} is the element at row p and column v in the matrix $\boldsymbol{M}$; m^{vp} is the element at row v and column p in the matrix $\boldsymbol{M}$; m^{pp} is the p^{th} diagonal element in $\boldsymbol{M}$; x^p and x^q are the p^{th} and q^{th} features of the input vector $\boldsymbol{X}$, respectively; $x_{\mathrm{train},i}^p$ is the p^{th} feature of the i^{th} sample in the training set; $x_{\mathrm{train},i}^v$ is the v^{th} feature of the i^{th} sample in the training set; $C_{a,i}^p$ is the coefficient of the term x^p; and $C_{b,i}^p$ is the constant term. Formula (25) represents the power flow distance between the i^{th} sample in the training set and the j^{th} sample in the testing set. It can be observed from (25) that the power flow distance between a testing sample and any training sample is the sum of the square terms, cross multiplication terms, linear terms, and constant terms of the input power flow $\boldsymbol{X}$ under this training sample.

Further analysis shows that:

1) The coefficients of the square terms and the cross multiplication terms in the distance expression are only related to elements of matrix $\boldsymbol{M}$, while the linear and constant terms are related to both the elements of $\boldsymbol{M}$ and the training sample.

2) Since the square term and cross multiplication term in $D_{\mathrm{M}}^2\left(\boldsymbol{X}_i^{\mathrm{train}}, \boldsymbol{X}\right)$ exist in both the numerator and denominator of the CCT estimator given by (25), they can be crossed out at the same time, therefore their values actually have no effect on the estimation result $\hat{y}_{\mathrm{CCT}}$.

3) The existence of the square terms and cross multiplication terms makes the distance meet the range requirement as $D_{\mathrm{M}}\left(\boldsymbol{X}_i^{\mathrm{train}}, \boldsymbol{X}\right) \geq 0$. If the square terms and cross multiplication terms are deleted, the remaining part will not meet the definition of distance and it is difficult to determine its range for linearization. Therefore, they are retained in the modification process of preventive control.

Since the variables in the TSA model involve data from two dimensions of input features and training samples, the stability constraints directly constituted by the trained TSA model without modification will have high complexity and the compute scale of the OPC problem will be very large. Taking the IEEE 10-machine 39-bus (10M39B) system as an example, assuming that 300 variables are selected as input features, the element amount of matrix $\boldsymbol{M}$ obtained after training can reach 90000, and the number of training samples usually needs to be more than 10000. Therefore, the specific formula of the TSA model is actually very complex and such a large-scale optimization model of preventive control with other constraints cannot be solved directly.

In order to solve the aforementioned complexity problem, the TSA model needs to go through the processes of key feature extraction and dominant sample selection to form a refined expression of transient stability constraints. The obtained expression needs to be further linearized as explained in Section III-E to be applied to the preventive control model as the final transient stability constraints.

C. Key Feature Extraction

It is found that the absolute values of most elements in matrix $\boldsymbol{M}$ are close to 0 [46], which makes it possible to realize feature extraction.

In essence, the calculation of the Mahalanobis distance is to rotate and stretch the two vectors $\boldsymbol{x}$ and $\boldsymbol{x}_i$ involved in the distance formula (17) successively. Since the property of distance requires $D_{\mathrm{M}}\left(\boldsymbol{X}_i^{\mathrm{train}}, \boldsymbol{X}\right) \geq 0$, it is necessary to ensure that matrix $\boldsymbol{M}$ is positive semidefinite and symmetric.

When the absolute values of some elements in $\boldsymbol{M}$ are close to 0, the corresponding terms in (25) are also approximately equal to 0. Therefore, in the presence of elements with absolute values significantly greater than 0, the deletion of terms about 0 has little effect on the value of distance theoretically.

In addition, the value of the coefficient γ in the MD-kernel regression given by (18) is mostly tens or even hundreds obtained by training, which enlarges the value of each term in distance $D_{\mathrm{M}}^2\left(\boldsymbol{X}_i^{\mathrm{train}},\boldsymbol{X}\right)$. The exp(·) function will also enlarge the value of the MD-kernel regression. The double amplification of coefficient γ and exp(·) function makes the absolute value difference of the elements in $\boldsymbol{M}$ be magnified by several orders of magnitude. Therefore, the accuracy of the stability estimation results will not be significantly reduced by simplifying the elements of $\boldsymbol{M}$ theoretically.

According to (25), since the features related to the undeleted elements are retained in the final stability constraint, the deletion of elements in matrix $\boldsymbol{M}$ can actually realize the extraction of key features. The way to realize the element selection of matrix $\boldsymbol{M}$ is to set the threshold of absolute value m_{thre}, which needs to be determined by balancing the requirements of estimation accuracy and the amounts of key features reserved.

The specific process of key feature extraction can be divided into five steps.

Step 1: define the initial threshold m_{thre} as 1/2 of the maximum absolute value of elements in matrix $\boldsymbol{M}$, i. e., $m_{\mathrm{thre}}=\max\left(|\boldsymbol{M}|\right)/2$.

Step 2: set the elements in matrix $\boldsymbol{M}$ with absolute value less than m_{thre} to be 0, and thus get the final distance metric matrix $\boldsymbol{M}_{\mathrm{fea}}$.

Step 3: use $\boldsymbol{M}_{\mathrm{fea}}$ to calculate the predicted CCT value of each sample in the testing set, and obtain the average prediction accuracy P_{fea} of the testing set when $\boldsymbol{M}_{\mathrm{fea}}$ is adopted.

Step 4: compare P_{fea} with the initial average accuracy P_0 of the testing set when using the matrix $\boldsymbol{M}$ for CCT prediction, and calculate the accuracy loss $\Delta P=P_0-P_{\mathrm{fea}}$. If $\Delta P\leq 0.05\%$, the matrix $\boldsymbol{M}_{\mathrm{fea}}$ at this time is the final distance metric matrix which can meet the requirement of key feature extraction. Otherwise, if $\Delta P>0.05\%$, set $\boldsymbol{M}_{\mathrm{fea}}=\boldsymbol{M}_{\mathrm{fea}}/2$ and return to *Step 2*.

Step 5: for the final distance metric matrix $\boldsymbol{M}_{\mathrm{fea}}$, from $i=1$ to $i=N_{\mathrm{fea}}$, judge whether the elements on the i^{th} row and i^{th} column are all 0. If so, it means that the feature $\boldsymbol{X}_i$ corresponding to the serial number i is redundant and needs to be removed from the initial feature set $\boldsymbol{X}$. After filtering, the remaining features constitute the key feature set $\boldsymbol{X}_{\mathrm{key}}$.

In the feature extracting process, the features which are treated as control variables in preventive control should be retained as much as possible to preserve the influence of those control variables on the stability assessment result. In this paper, the active power outputs of all generators are retained in the transient stability constraint after the feature extraction based on threshold m_{thre}. Therefore, the modified distance metric matrix $\boldsymbol{M}_{\mathrm{fea}}$ is finally obtained.

D. Dominant Sample Selection

After feature extraction with element screening of matrix $\boldsymbol{M}$ as the main means, the complexity problem from input features of the stability assessment model is solved to a certain extent. However, the solution of estimated CCT also needs the weighted sum of CCT values of all the training samples, which makes the model still have high complexity. Therefore, dominant sample selection is also an important modification for the trained CCT estimator to construct the transient stability constraints.

The basic idea of the TSA model based on the MD-kernel regression is to obtain the estimation result by using the samples most similar to the scenario to be estimated. The closer the power flow distance between the training sample and the scenario to be estimated, the higher the similarity between them, and the higher the weight of the CCT value of this training sample in the estimation model. Moreover, due to the double amplification effect of the coefficient $-\gamma/2$ and the exp(·) function in (18), the weights of training samples with different similarity degrees to the scenario to be estimated are quite different. Since the essence of the distance term D_{M}^2 is to measure the degree of similarity, the set of training samples that really determine the estimation result is the training samples with the smallest distance from the scenario to be estimated, which can be called as the dominant samples of stability assessment.

Inspired by the above analysis, it is noticed that seeking the dominant samples of preventive control is the key to construct the stability constraint. The process of preventive control is: given an initial power flow scenario $\boldsymbol{X}_0$, a new operation scenario $\boldsymbol{X}_{\mathrm{stable}}$ is obtained by changing the control variables $\boldsymbol{X}_u$ in $\boldsymbol{X}_0$, so as to meet the requirements of transient stability margin and other operation constraints. Meanwhile, the dependent variables $\boldsymbol{X}_d$ in the initial scenario $\boldsymbol{X}_0$ will change along with the changes of $\boldsymbol{X}_u$, so only part of features $\boldsymbol{X}_c$ in $\boldsymbol{X}_0$ remains unchanged before and after control. Since the change of load is generally not considered in the preventive control, $\boldsymbol{X}_c$ often refers to the active and reactive power consumptions of each load in the power system. Therefore, the load distances between the initial operation scenario and all training samples can be calculated, and the dominant samples for control can be obtained by selecting specified number of training samples with the smallest load distances.

Based on the above analysis, the specific procedure of the dominant sample selection method proposed in this paper contains the following four steps.

Step 1: define the features except $\boldsymbol{X}_c$ in the input features of all the training samples as the values of corresponding features in the initial power flow scenario $\boldsymbol{X}_0$, which means only the loads of all training samples are different from the initial power flow scenario $\boldsymbol{X}_0$, and the modified training set $\boldsymbol{X}_{\mathrm{mo}}^{\mathrm{TRAIN}}$ composed of the changed training samples is thus obtained. Calculate the load distance $D_{M(i_{mo},0)}^2$ between each sample in $\boldsymbol{X}_{\mathrm{mo}}^{\mathrm{TRAIN}}$ and the initial scenario $\boldsymbol{X}_0$, and select N_{load} samples with the smallest distance to constitute the dominant sample set $\boldsymbol{X}_1^{\mathrm{TRAIN}}$.

Step 2: calculate the power flow distance $D_{M(i,0)}^2$ between each sample in the original training sample set $\boldsymbol{X}^{\mathrm{TRAIN}}$ and the initial scenario $\boldsymbol{X}_0$, and select N_{add} samples with the smallest distance to constitute the dominant sample set $\boldsymbol{X}_2^{\mathrm{TRAIN}}$.

Step 3: select N_{thre} training samples whose CCT value is close to the threshold y_{thre} to form the dominant sample set $\boldsymbol{X}_3^{\mathrm{TRAIN}}$.

Step 4: take the union set of the three sample sets $\boldsymbol{X}_1^{\mathrm{TRAIN}}$,

$\boldsymbol{X}_2^{\text{TRAIN}}$ and $\boldsymbol{X}_3^{\text{TRAIN}}$ to get the final dominant sample set $\boldsymbol{X}_{\text{selected}}^{\text{TRAIN}}$ which contains N_{sel} training samples.

Therefore, tens of thousands of training samples in the TSA model can be reduced to any number of samples in the transient stability constraints of preventive control model according to the requirements of solving efficiency and estimation accuracy, so as to simplify the model scale of preventive control and eliminate the complexity problem caused by the large-scale training sample set.

E. Bi-level Linearization

After key feature extraction and dominant sample selection, the expression of stability constraint is obtained as:

$$D_M^2\left(\boldsymbol{X}_i^{\text{train}},\boldsymbol{X}\right)=\left(\boldsymbol{X}-\boldsymbol{X}_i^{\text{train}}\right)\boldsymbol{M}_{\text{fea}}\left(\boldsymbol{X}-\boldsymbol{X}_i^{\text{train}}\right)^{\text{T}} \tag{28}$$

$$\kappa_{\text{MD}}\left(\boldsymbol{X}_i^{\text{train}},\boldsymbol{X}\right)=\exp\left(-\frac{\gamma}{2}D_M^2\left(\boldsymbol{X}_i^{\text{train}},\boldsymbol{X}\right)\right) \tag{29}$$

$$\hat{y}_{\text{CCT}}\left(\boldsymbol{X}\right)=\frac{\sum_{i=1}^{N_{\text{sel}}}\kappa_{\text{MD}}\left(\boldsymbol{X}_i^{\text{train}},\boldsymbol{X}\right)y_i^{\text{train}}}{\sum_{i=1}^{N_{\text{sel}}}\kappa_{\text{MD}}\left(\boldsymbol{X}_i^{\text{train}},\boldsymbol{X}\right)}\geq y_{\text{thre}} \tag{30}$$

With the non-linear term D_M^2 and the exp(·) function in the constraint, it is impossible to solve the OPC model directly. Therefore, the linearization method is adopted to obtain the final solvable transient stability constraints. Since the terms in D_M^2 and the exp(·) function are nested in the MD-kernel regression as in (29), the linearization process is called bi-level linearization.

For the square terms and cross multiplication terms in the inner expression D_M^2 given by (25), the linearization method around the initial operation point is adopted:

$$\left(x^p\right)^2=2x_0^px^p-\left(x_0^p\right)^2 \tag{31}$$

$$x^px^q=x_0^px^q+x_0^qx^p-x_0^px_0^q \tag{32}$$

where x_0^p and x_0^q are the p^{th} and q^{th} features of the original operation scenario $\boldsymbol{X}_0$, respectively.

For the outer exp(·) function, the piecewise linearization method is adopted as:

$$D_i=\sum_{l=1}^{L_{\text{exp}}}\alpha_i^l\left(\overline{D}_i^l-\underline{D}_i^l\right)+\sum_{l=1}^{L_{\text{exp}}}\underline{D}_i^l\beta_i^l\quad\forall i\in N_{\text{sel}} \tag{33}$$

$$Z_i=\sum_{l=1}^{L_{\text{exp}}}\alpha_i^l\left(\overline{Z}_i^l-\underline{Z}_i^l\right)+\sum_{l=1}^{L_{\text{exp}}}\underline{Z}_i^l\beta_i^l\quad\forall i\in N_{\text{sel}} \tag{34}$$

$$\begin{cases}\overline{Z}_i^l=\exp\left(-\dfrac{\gamma}{2}\overline{D}_i^l\right)\\ \underline{Z}_i^l=\exp\left(-\dfrac{\gamma}{2}\underline{D}_i^l\right)\end{cases}\quad\forall l\in L_{\text{exp}},i\in N_{\text{sel}} \tag{35}$$

$$\alpha_i^l\leq\beta_i^l\quad\forall l\in L_{\text{exp}},i\in N_{\text{sel}} \tag{36}$$

$$\sum_{l=1}^{L_{\text{exp}}}\beta_i^l=1\quad\forall i\in N_{\text{sel}} \tag{37}$$

where D_i represents the $D_M^2\left(\boldsymbol{X}_i^{\text{train}},\boldsymbol{X}\right)$ from the initial scenario to the i^{th} training sample; Z_i represents the corresponding linearized MD-kernel regression; $\overline{D}_i^l$ and $\underline{D}_i^l$ are the upper and lower limits of the l^{th} segment of D_i, respectively; $\overline{Z}_i^l$ and $\underline{Z}_i^l$ are the linearized MD-kernel regression Z_i corresponding to $\overline{D}_i^l$ and $\underline{D}_i^l$, respectively; L_{exp} is the number of linearized segments; and α_i^l and β_i^l are two related variables introduced in the l^{th} segment of D_i for linearization, in which α_i^l is a continuous variable with a value ranging from 0 to 1 and β_i^l is a binary variable. The constraints (36) and (37) can guarantee that the values of D_i and Z_i can only be located in a single position within a certain segment.

Therefrom, the linearization of all non-linear terms is realized. By passing through the above three links of modifications, the final transient stability constraint can be concluded as the combination of (28), (31)-(38).

$$\sum_{i=1}^{N_{\text{sel}}}Z_iy_i^{\text{train}}\geq y_{\text{thre}}\sum_{i=1}^{N_{\text{sel}}}Z_i \tag{38}$$

Formula (38) is the linearized version of the stability constraint corresponding to (30) and the set of these final stability constraints is the detailed form of (15).

IV. Procedure of Data-driven TSA and OPC

In order to realize the direct solution of the OPC problem, it is also necessary to linearize the general power flow equality constraints shown as (9) and (10). This paper adopts the hot-start linear power flow equations proposed in [44] to supplant the general power flow equality constraints, which will lead to the change of some input variables in TSA model and the corresponding variables in preventive control.

After obtaining the linear analytical expression of transient stability constraints and power flow equality constraints, the overall procedure of data-driven TSA and OPC can be summarized for practical application.

A. Hot-start Linear Power Flow Equations

The non-linearity of the power flow equations leads to the non-convex of OPC problem, making the convergence or the global optimality not be guaranteed by the existing algorithms. The hot-start linear form of power flow equations proposed in [44] is introduced into the OPC model to remove the non-linear terms of power flow equations while the precision loss is small enough.

Through a series of variable transformations proposed in [44], the hot-start linear power flow equations can be expressed as:

$$P_{ij}=g_{ij}V_i^2-g_{ij}^P\frac{V_i^2+V_j^2}{2}-b_{ij}^P(\theta_{ij}-\theta_{ij,0})+g_{ij}^P\frac{V_{ij,L}^s}{2} \tag{39}$$

$$Q_{ij}=-b_{ij}V_i^2+b_{ij}^Q\frac{V_i^2+V_j^2}{2}-g_{ij}^Q(\theta_{ij}-\theta_{ij,0})-b_{ij}^Q\frac{V_{ij,L}^s}{2} \tag{40}$$

$$V_{ij,L}^s=2\frac{V_{i,0}-V_{j,0}}{V_{i,0}+V_{j,0}}\left(V_i^2-V_j^2\right)-V_{ij,0}^2 \tag{41}$$

$$g_{ij}^P=\left(g_{ij}c_{ij}^0+b_{ij}s_{ij}^0\right)+\left(g_{ij}c_{ij}^1+b_{ij}s_{ij}^1\right)\theta_{ij,0}=g_{ij}\cos\theta_{ij,0}+b_{ij}\sin\theta_{ij,0} \tag{42}$$

$$b_{ij}^{P}=\left(g_{ij}c_{ij}^{1}+b_{ij}s_{ij}^{1}\right)V_{i,0}V_{j,0} \tag{43}$$

$$b_{ij}^{Q}=\left(-g_{ij}s_{ij}^{0}+b_{ij}c_{ij}^{0}\right)-\left(g_{ij}s_{ij}^{1}-b_{ij}c_{ij}^{1}\right)\theta_{ij,0}=-g_{ij}\sin\theta_{ij,0}+b_{ij}\cos\theta_{ij,0} \tag{44}$$

$$g_{ij}^{Q}=\left(g_{ij}s_{ij}^{1}-b_{ij}c_{ij}^{1}\right)V_{i,0}V_{j,0} \tag{45}$$

$$\begin{cases}s_{ij}^{1}=\cos\theta_{ij,0}\\ s_{ij}^{0}=\sin\theta_{ij,0}-\theta_{ij,0}\cos\theta_{ij,0}\end{cases} \tag{46}$$

$$\begin{cases}c_{ij}^{1}=-\sin\theta_{ij,0}\\ c_{ij,0}^{0}=\cos\theta_{ij,0}+\theta_{ij,0}\sin\theta_{ij,0}\end{cases} \tag{47}$$

where $\theta_{ij,0}$, $V_{i,0}$, and $V_{j,0}$ are the initial values of the variables θ_{ij}, V_i, and V_j, respectively.

The formulas (39) and (40) are the final power flow equality constraints adopted in the OPC model proposed in this paper. Since it can be regarded as a linear equation only when the square of voltage amplitude V^2 is regarded as an independent variable, all the voltage amplitudes V in the OPC model including the transient stability constraints should be replaced by V^2 to ensure the consistency of variables.

B. Variables of TSA Model and OPC Model

Since the transient stability constraints in the OPC model are composed of the modified data-driven TSA model, the input variables of TSA model should be consistent with the OPC model. Therefore, the input variables of the data-driven TSA model, which are actually the variables in the pre-fault power flow vector $\boldsymbol{X}$ of the mapping (16), are listed in Table I.

TABLE I
VARIABLES IN PRE-FAULT POWER FLOW VECTOR

Element type	Variable	Description
Bus	V_a^2, θ_a	Square of voltage amplitude and phase angle of bus a
AC line (including transformer)	P_{ab}, Q_{ab}	Active and reactive power injected from bus a through line L_{ab}
Generator	$P_{G,a}$, $Q_{G,a}$	Active and reactive power generations of the generator connected to bus a
Load	$P_{L,a}$, $Q_{L,a}$	Active and reactive power consumptions of the load connected to bus a

Compared with the variables listed in [42], there are two differences in the variables describing pre-fault power flow.

1) In order to meet the requirements of the hot-start linear power flow equality constraints introduced in Section IV-A, the square of voltage amplitude V_a^2 and phase angle θ_a are taken as variables of each bus.

2) Since the capacitive charging power Q_{ca} of each AC line or each transformer is proportional to V_a^2 of the corresponding bus, the variable Q_{ca} is not included in the input feature set so as to avoid the collinearity problem of the data-driven model.

The selected variables are finally normalized to the range [−1, 1] through the method employed in [42].

It should be noted that the variable space composed of control variables and dependent variables in the OPC model is different from that of input variables in the TSA model. There is an intersection between the two variable spaces. Specifically, the load related variables $P_{L,a}$ and $Q_{L,a}$ are constant in the preventive control model. And there maybe also some variables in the control model that are not considered in the TSA model, most of which are the redundant features eliminated in the feature extraction process of data-driven method.

C. Overall Procedure

The overall procedure of the proposed data-driven TSA and OPC can be illustrated as Fig. 1. The two green dashed boxes in this figure represent the construction processes of the data-driven transient stability constraint and the low non-linearity power flow equality constraint, respectively.

The central issue of this procedure is the construction of the OPC model, in which the analytical representation of data-driven transient stability constraint is the most crucial. Before that, a data-driven TSA model based on the non-parametric regression is built and well-trained to determine the transient stability of power system. After that, considering the requirements of preventive control, the trained TSA model is processed successively by key feature extraction, dominant sample selection, and bi-level linearization. The explicit expression of data-driven transient stability constraints is thus obtained.

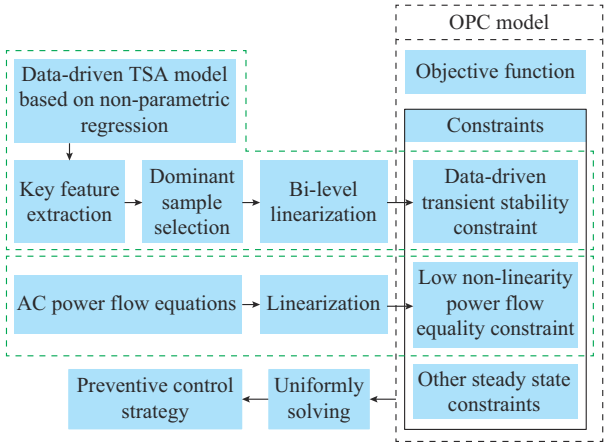

Fig. 1. Overall procedure of data-driven TSA and OPC.

In addition, the AC power flow equations are considered in the OPC model. In order to facilitate the efficient and unified solving of this model, a hot-start linear power flow model proposed in [44] is obtained by the linearization of the AC power flow equations and applied to the OPC model as the low non-linearity power flow equality constraint.

Therefore, the proposed OPC model is transformed into a MILP problem which can be solved without iteration, thus avoiding the non-convergence problem of the traditional iterative methods. And a higher quality solution is obtained by comprehensively considering the global constraints.

V. Case Study

The validity of the proposed data-driven TSA and OPC is tested and verified in the IEEE 10M39B system and IEEE 300-bus system. The PSASP 7.40, an integrated power system simulation tool, is applied to build the simulation models of the two systems and calculate the CCT value of different samples. The static parameters of two simulation systems are set referring to the models in MATPOWER 6.0 [47], and the dynamic parameters are taken from the real power system. All the simulations and calculations in this paper are conducted on an 8-core workstation, each with a 3.50 GHz CPU (Intel Xeon CPU E5-2637). The solver for MD-kernel regression training is Python 3.7 and the solver for optimal control of power system is CPLEX 12.4. The result analysis of IEEE 10M39B system will be introduced in the following three subsections and the results of IEEE 300-bus system will be presented in Section V-D.

A. Effect Analysis of One Single Case

The IEEE 10M39B system is constructed of 10 power units, 34 transmission lines, 12 transformers, and 21 loads. A case to be controlled with CCT value of 0.2719 in initial operation scenario is used to explain the effect of the proposed method in detail and more cases will be shown in Section V-B.

1) Data-driven TSA

In this paper, a sample is an operation scenario of the system labelled with the CCT value of a given fault. As shown in Table I, each sample has 8 kinds of features as the inputs and the CCT value as the output. After feature processing, 298 features are selected as the inputs of the data-driven TSA model. The generated 12400 samples are randomly divided into a training set consisting of 10000 samples and a testing set consisting of 2400 samples. The mean accuracy rate (MAR) of the testing set is 98.74%, which is defined as:

$$MAR = \frac{1}{N_{\text{test}}} \sum_{n=1}^{N_{\text{test}}} \left(1 - \left| \frac{\hat{y}_{\text{CCT},n} - y_{\text{CCT},n}}{y_{\text{CCT},n}} \right| \right) \times 100\% \tag{48}$$

where $\hat{y}_{\text{CCT},n}$ and $y_{\text{CCT},n}$ are the estimation of CCT and the actual value of CCT, respectively; and N_{test} is the total number of testing samples.

The accuracy of most cases in the testing set is higher than 96%, which indicates that the non-parametric regression model has high accuracy in TSA. And the estimation results can meet the requirement of conservatism in the power system according to [42].

2) Key Feature Extraction

Table II shows the accuracy of the proposed key feature extraction. It can be observed that the MAR of the testing set is still quite high even with only a few key features. The MAR of the testing set can reach 98.69%, only 0.05% lower than the MAR with full features, if 1297 elements of 88804 elements in the matrix $\boldsymbol{M}$ are extracted. Therefore, 197 redundant features are eliminated. The proposed key feature extraction can significantly reduce the scale of the problem while maintaining the high accuracy.

TABLE II
Accuracy of Key Feature Extraction

Threshold of absolute value of elements in $\boldsymbol{M}$	Number of remaining elements in $\boldsymbol{M}$	MAR of testing set (%)	MAR with full features (%)
0.052700	10	88.01	
0.026300	34	90.51	
0.013200	92	93.25	
0.006600	205	95.15	
0.003300	409	97.58	98.74
0.001600	639	98.39	
0.000820	889	98.47	
0.000410	1110	98.51	
0.000206	1297	98.69	

3) Dominant Sample Selection

There are 10000 samples in the training set. The weighting coefficients of the CCT values of the training samples on the estimated CCT value of the initial case in the data-driven TSA model are shown in Fig. 2, in which the training samples are sorted by the Mahalanobis distance between each training sample and the case to be controlled from small to large.

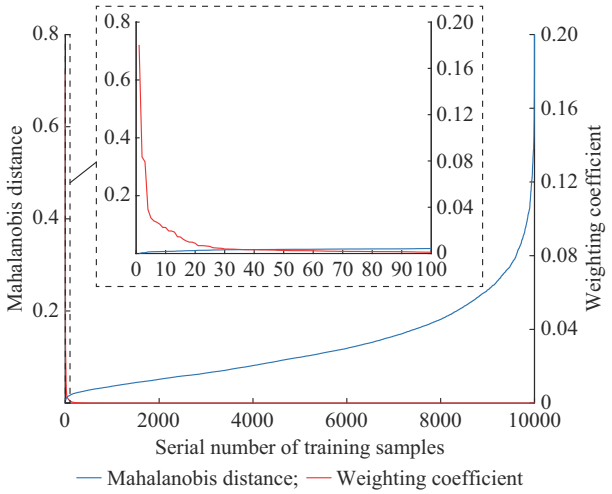

Fig. 2. Weighting coefficients of training samples in TSA model with different Mahalanobis distances.

The weighting coefficients in Fig. 2 are actually the parameter μ_i in (19) and (20). The closer the weighting coefficient is to 1, the greater the influence of the corresponding training sample on the estimated CCT value. On the contrary, when the weight is close to 0, the influence of the corresponding training sample on the estimated CCT value is negligible. Therefore, it can be observed from Fig. 2 that most of the training samples have little impact on the estimated result. Only a few samples with small Mahalanobis distance and large weighting coefficients have significant impact on the estimated result.

In the dominant sample selection process, 80% of the selected training samples have the smallest load distance from the case to be controlled and 20% of the selected samples

have the smallest power flow distance from the case to be controlled. The estimation accuracy of the data-driven TSA model with different numbers of retained training samples that are sorted by the distance from small to large is shown in Fig. 3, in which the dominant samples are a part of training samples with the smallest Mahalanobis distance from the case to be controlled.

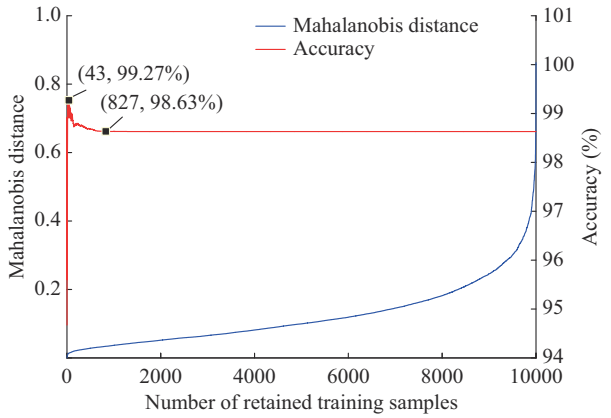

Fig. 3. Estimation accuracies of data-driven TSA model with different numbers of retained training samples.

It can be observed that the proposed dominant sample selection method has little loss of accuracy. Besides, abandoning the ineffective training samples can even improve the estimation accuracy of the initial operation point. Based on the analysis, only 50 samples are selected as dominant samples in the final model, which can already meet the accuracy requirement.

4) OPC Strategy

In the proposed OPC model, the MD-kernel function is linearized into 5 segments, which is shown as Fig. 4.

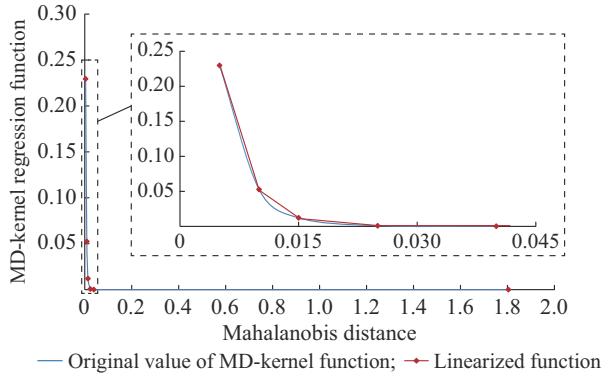

Fig. 4. Piecewise linearization of MD-kernel regression function.

Each segment corresponds to a binary variable. If the dominant sample selection is not adopted, there will be 50000 binary variables in the model, which makes the scale of the problem too large to be solved efficiently. After the dominant sample selection, only 50 samples remain, which significantly reduce the number of binary variables to 250 and thus greatly reduce the scale of the problem.

The CCT results after the preventive control are shown in Table III. CCT control targets are set separately. It can be observed that all the CCT control targets can be achieved with the proposed OPC model. The CCT value of time-domain simulation after control is also higher than the preset threshold of CCT value and is very close to the estimated CCT value, which indicates that the proposed OPC model can realize integrated optimization and control of the power system.

TABLE III
CCT Results After Control

Threshold of CCT value	Estimated CCT value of proposed OPC model	CCT value of time-domain simulation after control	Computation time (s)
0.30	0.3109	0.3218	2.06
0.35	0.3596	0.3625	9.20
0.40	0.4048	0.4125	4.63
0.42	0.4249	0.4344	2.52

After control, the sample with the smallest Mahalanobis distance to the case to be controlled in the entire training set is already within the selected 50 samples obtained from the dominant sample selection process, which indicates that the proposed dominant sample selection method can select the appropriate samples.

The computation time of the proposed OPC model is different with various threshold settings, which is mainly determined by the density of dominant samples near the threshold. In general, the proposed OPC model is of high computational efficiency compared with other existing control methods and can be applied to large-scale power system.

B. Different Operation Conditions

Various operation conditions to be controlled with different CCT values are further tested with the proposed OPC model, and the results are shown in Table IV.

TABLE IV
Control Effect Under Different Operation Conditions

CCT of initial operation condition	Threshold of CCT value	Estimated CCT value of proposed OPC model	CCT value of time-domain simulation after control	Computation time (s)
0.1938	0.4500	0.4560	0.4641	47.22
0.2125	0.4500	0.4597	0.4615	37.93
0.2531	0.4500	0.4659	0.4734	13.16
0.3219	0.4500	0.4511	0.4547	6.25

It can be observed that the proposed model is effective and efficient under different operation conditions. In fact, dozens of initial operation conditions with different CCT values have been tested, and the results fully demonstrate the robustness of the proposed OPC model.

Furthermore, as for the case with initial CCT value of 0.2531, the CCT values under different settings of the data-driven stability constraints after preventive control are analyzed as shown in Table V. In Table V, the difference between ④ and ③ is the error caused by the bi-linearization process, which is about 0.76%; the difference between ⑤

and ⑦ is the error caused by the dominant sample selection process, which is about 1.64%; the difference between ⑥ and ⑦ is the error caused by the key feature extraction process, which is about 0.96%. It can be observed that the errors in each step during the construction of the data-driven OPC model are relatively small. These kinds of errors are also calculated in sufficient cases, and the results show that these errors are acceptable.

TABLE V
ERRORS OF EACH STEP IN CONTROL PROCESS

Item	CCT value	Serial number
CCT value under initial operation condition	0.2531	①
Threshold of CCT value	0.4500	②
CCT value after control obtained by analytical stability constraint	0.4500	③
CCT value after control (dominant sample set, modified matrix $\boldsymbol{M}_{\text{fea}}$)	0.4534	④
CCT value after control (dominant sample set, initial matrix $\boldsymbol{M}$)	0.4585	⑤
CCT value after control (full sample set, modified matrix $\boldsymbol{M}_{\text{fea}}$)	0.4616	⑥
CCT value after control (full sample set, initial matrix $\boldsymbol{M}$)	0.4659	⑦

Considering the conservatism of the control results, the total error of the control results is negative, i.e., the real CCT value of the operation scenario after control needs to be greater or equal to the estimated value obtained by the proposed OPC model. In the TSA process, the data-driven TSA method based on kernel regression can obtain the estimation results under the required degree of conservatism [42], so as to control the direction of estimation error. And in the OPC process, the conservatism can be improved by properly adding the training samples with lower CCT value into the dominant sample set. The control of the error will be introduced in the further research.

C. Comparison with Classical Method

The comparison with a classical method is necessary to illustrate the control performance. The most common classical method is to iteratively solve the operation condition that meets the stability requirements. Generally, the intelligent optimization algorithm is used to search the feasible point, and the stability judgment needs to call time-domain simulation repeatedly. Therefore, as a typical classical method, the combination of PSO algorithm and time-domain simulation [48] is used to solve the preventive control problem, and the results are compared with the proposed method.

The fitness function in the PSO algorithm is constructed by the regulation of generator output and the stability constraint in the form of penalty function, which can be expressed as:

$$f_{\text{fit}} = w_1 \sum_i \left| P_{G,i} - P_{G,i}^0 \right| - w_2 \left(y_{\text{CCT}} - y_{\text{thre}} \right) \tag{49}$$

where y_{CCT} is the CCT value of the current operation condition calculated by time-domain simulation; and w_1 and w_2 are the penalty coefficients of generator output regulation and stability constraint, respectively. The position vector of particle i is constructed by the output of each generator, which can be represented as $\boldsymbol{X}_i = [P_{G,1}, P_{G,2}, \ldots, P_{G,N_{gen}}]$. The dimension, size, iterations, and the maximum speed of particle are set to be 10, 20, 1000, and 0.5, respectively.

The same operation condition with CCT value of 0.2531 in Table IV is selected as the comparative case, and the results are shown as Table VI. To ensure the consistency, the time-domain simulation is implemented by calling the PSASP 7.40 in the iterative process.

TABLE VI
CONTROL RESULTS OF DIFFERENT METHODS

Method	Estimated CCT value of proposed OPC model	CCT value of time-domain simulation after control	Absolute value of total regulation	Computation time (s)
Proposed method	0.4659	0.4734	14.4980	13.16
Classical method		0.4673	13.4702	1228.00

From the control results shown in Table VI, it can be observed that the CCT value of the proposed method is slightly higher than that of the classical method, mainly due to the consideration of conservatism in the proposed method. Moreover, although the total output regulation of the proposed method is slightly higher than the classical method, the calculation time of the former is much shorter than that of the latter. These results show that the control effect of the proposed method is not inferior to that of the classical method, and can greatly improve the computational efficiency.

D. Preventive Control in Larger System

The proposed method is further verified in the IEEE 300-bus system, which consists of 69 generators, 206 loads, and 411 transmission lines. The initial pre-fault power flow vector obtained from PSASP contains 3616 variables. After data preprocessing and feature selection [49], 2535 variables are retained as the inputs of the data-driven TSA model. The training set and testing set of a given contingency contain 32000 and 8000 samples, respectively. The MAR of the testing set in IEEE 300-bus system is 97.16%.

After passing through the key feature extraction process, 352 key features are extracted, which have a greater impact on the CCT value and thus the control effects caused by them are more obvious.

The control effects of three initial operation conditions with different thresholds in the IEEE 300-bus system are shown in Table VII, which indicates that the proposed method is applicable to the stability assessment and preventive control of a larger power system, and further demonstrates its application prospects in the actual power grid. For each operation conditions to be controlled, 100 dominant samples are selected. Note that the effective application of the proposed method in actual system is based on the sufficient reserve of training samples and the construction of TSA model with high prediction accuracy, which should be realized before the implementation of preventive control process.

It is worth noting that the computation time does not increase significantly with the increase of system scale. This is because the number of binary variables added by the linearization is the main factor that affects the computation time of the optimization model, and the process of dominant sample selection can greatly decrease the number of binary variables. Since the number of dominant samples is controllable, the computation efficiency is thus guaranteed, which is one of the most significant advantages compared with other methods.

TABLE VII
CONTROL EFFECT OF THREE INITIAL OPERATION CONDITIONS WITH DIFFERENT THRESHOLDS IN IEEE 300-BUS SYSTEM

Initial operation condition	CCT of initial operation condition	Threshold of CCT value	Estimated CCT value of proposed model	CCT value of time-domain simulation after control	Computation time (s)
①	0.1938	0.3000	0.3102	0.3234	140.79
①	0.1938	0.4000	0.4094	0.4141	252.77
②	0.2844	0.3500	0.3690	0.3844	98.42
②	0.2844	0.4500	0.4571	0.4734	116.50
③	0.3438	0.5000	0.5015	0.5047	75.71
③	0.3438	0.6000	0.6167	0.6344	92.63

VI. CONCLUSION

An analytical representation method of data-driven transient stability constraint in the OPC model is proposed in this paper. It is constructed from a TSA model based on nonparametric regression. After the kernel regression model is built for TSA, key feature extraction and dominant sample selection are proposed, which can significantly reduce the scale of the problem while maintaining high accuracy of TSA. After linearizing the transient stability constraints and the AC power flow equality constraints, the optimal control problem is finally formulated as an MILP problem, which can be solved without iterations. Case studies in the IEEE 10M39B system and IEEE 300-bus system show that the proposed method can greatly reduce the scale of the preventive control model while maintaining the key features of TSA, and the requirements of high accuracy and high efficiency are met at the same time. All the cases can meet the transient stability constraints after the preventive control, and the computation time indicates that the proposed method has broad application prospects.

It is worth noting that the proposed method is a general method to solve the optimal power flow problem with data-driven dynamic security constraints. Future work will focus on the effectiveness of this method applied in actual systems, changed topologies, and other stability problems.

REFERENCES

[1] S. Wu, L. Zheng, W. Hu *et al.*, "Improved deep belief network and model interpretation method for power system transient stability assessment," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 1, pp. 27-37, Jan. 2020.

[2] W. Hu, Z. Lu, S. Wu *et al.*, "Real-time transient stability assessment in power system based on improved SVM," *Journal of Modern Power Systems and Clean Energy*, vol. 7, no. 1, pp. 26-37, Jan. 2019.

[3] J. Lv and M. Pawlak, "Additive modeling and prediction of transient stability boundary in large-scale power systems using the group LASSO algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 113, pp. 963-970, Dec. 2019.

[4] F. Tian, X. Zhou, and Z. Yu, "Optimization method of transient stability preventive control based on sensitivity analysis and time domain simulation," *Electric Power Automation Equipment*, vol. 38, no. 7, pp. 155-61, Jul. 2018.

[5] Y. Zhou, J. Wu, L. Ji *et al.*, "Transient stability preventive control of power systems using chaotic particle swarm optimization combined with two-stage support vector machine," *Electric Power Systems Research*, vol. 155, pp. 111-120, Feb. 2018.

[6] C. Liu, K. Sun, Z. H. Rather *et al.*, "A systematic approach for dynamic security assessment and the corresponding preventive control scheme based on decision trees," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 717-730, Mar. 2014.

[7] S. Abhyankar, G. Geng, M. Anitescu *et al.*, "Solution techniques for transient stability-constrained optimal power flow–part I," *IET Generation, Transmission & Distribution*, vol. 11, no. 12, pp. 3177-3185, Aug. 2017.

[8] D. Gan, R. J. Thomas, and R. D. Zimmerman, "Stability-constrained optimal power flow," *IEEE Transactions on Power Systems*, vol. 15, no. 2, pp. 535-540, May 2000.

[9] J. Zhu, *Optimization of Power System Operation*. Estonia: IEEE and Wiley Press, 2009.

[10] Y. Xu, Z. Y. Dong, Z. Xu *et al*., "Power system transient stability-constrained optimal power flow: a comprehensive review," in *Proceedings of PES General Meeting*, San Diego, USA, Jul. 2012, pp. 1-7.

[11] S. Abhyankar, G. Geng, M. Anitescu *et al.*, "Solution techniques for transient stability-constrained optimal power flow–part II," *IET Generation, Transmission & Distribution*, vol. 11, no. 12, pp. 3186-3193, Aug. 2017.

[12] L. T. Biegler, A. M. Cervantes, and A. Wächter, "Advances in simultaneous strategies for dynamic process optimization," *Chemical Engineering Science*, vol. 57, no. 4, pp. 575-593, Feb. 2002.

[13] P. L. C. Wieler, R. Kuiava, and W. F. S. Souza, "Transient stability constrained optimal power flow based on trajectory sensitivity for power dispatch of distributed synchronous generators," *IEEE Latin America Transactions*, vol. 18, no. 7, pp. 1247-1254, May 2000.

[14] D. B. Leineweber, I. Bauer, H. G. Bock *et al.*, "An efficient multiple shooting based reduced SQP strategy for large-scale dynamic process optimization–part 1: theoretical aspects," *Computers & Chemical Engineering*, vol. 27, no. 2, pp. 157-166, Feb. 2003.

[15] Q. Jiang and Z. Huang, "An enhanced numerical discretization method for transient stability constrained optimal power flow," *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1790-1797, Nov. 2010.

[16] Q. Jiang and G. Geng, "A reduced-space interior point method for transient stability constrained optimal power flow," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1232-1240, Feb. 2010.

[17] I. A. Hiskens and M. A. Pai, "Power system applications of trajectory sensitivities," in *Proceedings of 2002 IEEE PES Winter Meeting*, New York, USA, Jan. 2002, pp. 1200-1205.

[18] H. Zhang, S. Abhyankar, E. Constantinescu *et al.*, "Discrete adjoint sensitivity analysis of hybrid dynamical systems with switching," *IEEE Transactions on Circuits & Systems I Regular Papers*, vol. 64, no. 5, pp. 1247-1259, Jan. 2017.

[19] G. Geng, V. Ajjarapu, and Q. Jiang, "A hybrid dynamic optimization approach for stability constrained optimal power flow," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2138-2149, Sept. 2014.

[20] A. Pizano-Martinez, C. R. Fuerte-Esquivel, and D. Ruiz-Vega. "A new

practical approach to transient stability-constrained optimal power flow," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1686-1696, Aug. 2011.
[21] N. Mo, Z. Zou, K. Chan *et al*., "Transient stability constrained optimal power flow using particle swarm optimization," *IET Generation, Transmission & Distribution*, vol. 1, no. 3, pp. 476-483, May 2007.
[22] C. Ye and M. Huang, "Multi-objective optimal power flow considering transient stability based on parallel NSGA-II," *IEEE Transactions on Power Systems*, vol. 30, no. 2, pp. 857-866, Mar. 2015.
[23] X. Zhang, R. Dunn, and F. Li, "Stability constrained optimal power flow for the balancing-market using genetic algorithms," in *Proceedings of PES General Meeting*, Tampa, USA, Jun. 2007, pp. 932-937.
[24] O. A. Alimi, K. Ouahada, and A. M. Abu-Mahfouz, "A review of machine learning approaches to power system security and stability," *IEEE Access*, vol. 8, pp. 113512-113531, Jun. 2020.
[25] Y. Xu, Z. Dong, J. Zhao *et al.*, "A reliable intelligent system for real-time dynamic security assessment of power systems," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1253-1263, Aug. 2012.
[26] P. Pavani and S. N. Singh, "Support vector machine based transient stability identification in distribution system with distributed generation," *Electric Power Components and Systems*, vol. 44, no. 1, pp. 60-71, Jan. 2016.
[27] T. Amraee and S. Ranjbar, "Transient instability prediction using decision tree technique," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3028-3037, Aug. 2013.
[28] N. A. Nguyen, H. A. Quyen, T. N. Le *et al.*, "An improvement forward floating search algorithm for feature selection in power system transient stability classification," in *Proceedings of AETA International Conference*, Ho Chi Minh City, Vietnam, Dec. 2015, pp. 167-174.
[29] L. Zhu, D. J. Hill, and C. Lu, "Hierarchical deep learning machine for power system online transient stability prediction," *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 2399-2411, May 2020.
[30] J. An, J. Yu, Z. Li *et al*., "A data-driven method for transient stability margin prediction based on security region," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 6, pp. 1060-1069, Dec. 2020.
[31] Y. Zhou, J. Wu, L. Ji *et al.*, "Two-stage support vector machines for transient stability prediction and preventive control of power systems," *Proceedings of the CSEE*, vol. 38, no. 1, pp. 137-147, Jan. 2018.
[32] F. Tian, X. Zhou, Z. Yu *et al.*, "A preventive transient stability control method based on support vector machine," *Electric Power Systems Research*, vol. 170, pp. 286-293, May 2019.
[33] I. Genc, R. Diao, V. Vittal *et al.*, "Decision trees-based preventive and corrective control applications for dynamic security enhancement in power systems," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1611-1619, Aug. 2010.
[34] H. Yuan and Y. Xu, "Preventive-corrective coordinated transient stability dispatch of power systems with uncertain wind power," *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 3616-3626, Sept. 2020.
[35] S. Bruno, E. D. Tuglie, and M. L. Scala, "Transient security dispatch for the concurrent optimization of plural postulated contingencies," *IEEE Transactions on Power Systems*, vol. 17, no. 3, pp. 707-714, Aug. 2002.
[36] T. Liu, Y. Liu, J. Liu *et al*., "A Bayesian learning based scheme for online dynamic security assessment and preventive control," *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 4088-4099, Sept. 2020.
[37] J. L. Cremer, I. Konstantelos, S. H. Tindemans *et al.*, "Data-driven power system operation: exploring the balance between cost and risk," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 791-801, Jan. 2019.
[38] Y. Xu, Z. Y. Dong, L. Guan *et al.*, "Preventive dynamic security control of power systems based on pattern discovery technique," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1236-1244, Aug. 2012.
[39] N. Mo, Z. Zou, K. W. Chan *et al.*, "Transient stability constrained optimal power flow using particle swarm optimization," *IET Generation, Transmission & Distribution*, vol. 1, no. 3, pp. 476-483, May 2007.
[40] Y. Jin, H. Wang, T. Chugh *et al.*, "Data-driven evolutionary optimization: an overview and case studies," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 3, pp. 442-458, Jun. 2018.
[41] H. Wang, Y. Jin, and J. O. Jansen, "Data-driven surrogate-assisted multiobjective evolutionary optimization of a trauma system," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 6, pp. 939-952, Dec. 2016.
[42] X. Liu, Y. Min, L. Chen *et al.*, "Data-driven transient stability assessment based on kernel regression and distance metric learning," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 1, pp. 27-36, Jan. 2021.
[43] T. Liu, Y. Liu, L. Xu *et al.*, "Non-parametric statistics-based predictor enabling online transient stability assessment," *IET Generation, Transmission & Distribution*, vol. 12, no. 21, pp. 5761-5769, Nov. 2018.
[44] Z. Yang, A. Bose, H. Zhong *et al.*, "Optimal reactive power dispatch with accurately modeled discrete control devices: a successive linear approximation approach," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2435-2444, May 2017.
[45] X. Liu, X. Zhang, L. Chen *et al*., "Data-driven transient stability assessment model considering network topology changes via Mahalanobis kernel regression and ensemble learning," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 6, pp. 1080-1091, Dec. 2020.
[46] S. Xiang, F. Nie, and C. Zhang, "Learning a Mahalanobis distance metric for data clustering and classification," *Pattern Recognition*, vol. 41, no. 12, pp. 3600-3612, Dec. 2008.
[47] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER: steady-state operations, planning and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12-19, Feb. 2011.
[48] S. Xia, K. W. Chan, X. Bai *et al.*, "Enhanced particle swarm optimisation applied for transient angle and voltage constrained discrete optimal power flow with flexible AC transmission system," *IET Generation, Transmission & Distribution*, vol. 9, no. 1, pp. 61-74, Jan. 2015.
[49] N. Spolaôr, M. C. Monard, G. Tsoumakas *et al*., "A systematic review of multi-label feature selection and a new method based on label construction," *Neurocomputing*, vol. 180, pp. 3-15, Mar. 2016

Yiwei Fu received the B.S. degree in electrical engineering from Tsinghua University, Beijing, China, in 2016, where she is currently pursuing the Ph.D. degree. Her research interests include power system stability and data-driven techniques.

Xiaohua Zhang received the B.S. and M.S. degrees in electrical engineering from Nanchang University, Nanchang, China, in 1984 and 2005, respectively. He is now a Professor Level Senior Engineer. He is currently working at the State Grid Jibei Electric Power Company Limited, Beijing, China. His research interests include operation and scheduling of large-scale power grid.

Lei Chen received the B. Sc. and Ph. D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2003 and 2008, respectively. Since 2008, he has been with the Department of Electrical Engineering, Tsinghua University, where he is now an Associate Professor. He was a recipient of the Excellent Young Scientists Fund of National Natural Science Foundation of China in 2019. His research interests include dynamic analysis and control of power systems.

Zengyao Tian received the B.S. degree in electrical engineering from Northeast Electric Power University, Jilin, China, in 2002. He is now working at Northeast Branch of State Grid Corporation of China, Shenyang, China. His research interests include power system operation and control.

Kaiyuan Hou received the B.S. M.S., and Ph.D. degrees in electrical engineering from Central China University of Science and Engineering, Hefei, China, Northeast Electric Power University, Jilin, China, and Tsinghua University, Beijing, China, in 1996, 2001, and 2005, respectively. He is now working at Northeast Branch of State Grid Corporation of China, Shenyang, China. His research interests include power grid dispatching operation, power grid security management, stability analysis and control.

Haikuan Wang received the B.S. and M.S. degrees in electrical engineering from Northeast Electric Power University, Jilin, China, in 1995 and 2007, respectively. He is now a Senior Engineer in Northeast Branch of State Grid Corporation of China, Shenyang, China. His research interests include power system operation and control.