

分类号: TM712

单位代码: 10335

密 级: 非涉密论文

学 号: 21710083

浙江大学

硕士学位论文



中文论文题目: 数据驱动的考虑温控负荷和风电不确定性的电力系统可靠性分析研究

英文论文题目: **Data-driven Power System Reliability Analysis Considering Uncertainty of TCLs and Wind Power**

申请人姓名: 庄欣然

指导教师: 丁一 教授

专业名称: 电气工程

研究方向: 电力系统可靠性分析

所在学院: 电气工程学院

论文提交日期 2020 年 1 月

数据驱动的考虑温控负荷和风电不确定性的
电力系统可靠性分析研究



论文作者签名:_____

指导教师签名:_____

论文评阅人 1: _____

评阅人 2: _____

评阅人 3: _____

答辩委员会主席: _____

委员 1: _____

委员 2: _____

委员 3: _____

委员 4: _____

答辩日期: _____

Data-driven Power System Reliability Analysis
Considering Uncertainty of TCLs and Wind Power



Author's signature:_____

Supervisor' s signature:_____

External Reviewers: _____

Examining Committee Chairperson:

Examining Committee Members:

Date of oral defence: _____

浙江大学研究生学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的
研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发
表或撰写过的研究成果，也不包含为获得 浙江大学 或其他教育机构的学位或
证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文
中作了明确的说明并表示谢意。

学位论文作者签名: 签字日期: 年 月 日

学位论文版权使用授权书

本学位论文作者完全了解 浙江大学 有权保留并向国家有关部门或机构送交本论文的复印件和磁盘，允许论文被查阅和借阅。本人授权 浙江大学 可以将学位论文的全部或部分内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

(保密的学位论文在解密后适用本授权书)

学位论文作者签名: 导师签名:

签字日期: 年 月 日 签字日期: 年 月 日

致 谢

时光荏苒，转眼间这已经是我在浙大校园的第七个年头。回望过往的点滴，我不禁感叹这段求学生涯早已在我的青春里留下了浓墨重彩的一笔。在这即将告别校园生活的时刻，我想向一路上为我提供支持和陪伴的师长和朋友们表示真诚的感谢。

本论文是在我的研究生导师丁一教授的悉心指导下完成的，特别感谢丁老师对我的帮助和鼓励。在我的研究生生涯中，丁老师在课题研究、科研项目、论文撰写与修改等各方各面都为我提供了无微不至的指导和关怀，在我遭遇困难挫折时为我耐心地答疑解惑。丁老师独到开阔的治学思路、严谨负责的工作态度，为我们树立了学习的标杆。在此我想向丁老师表示最真挚的感谢，您的教诲是我人生路上的巨大财富。

感谢叶承晋博士对我的指导和帮助，感谢您一直以来的积极引导和鼓励。感谢课题组的每一位同伴，你们的支持和陪伴为我的求学生涯增添了无数美好的色彩。感谢一路走来相识的每一位志同道合的朋友，你们是鼓舞我勇敢面对生活和学习工作中的挑战的强大动力。感谢每一位在我迷惘时为我排忧解难的朋友和前辈们，和你们的每一次交流都让我受益匪浅，你们的帮助让我对未来的人生有了更加积极的态度，

最后，我要衷心感谢我的父母，感谢你们多年来的养育之恩，感谢你们一直以来对我的默默支持，十分感恩能够拥有这样幸福美满的家庭，你们永远都会是我人生路上的坚实后盾。

庄欣然

二〇二〇年一月

于求是园

摘要

随着国家对节能减排、绿色环保的重视程度日益提高,大力发展风力发电等清洁能源已经成为能源领域的一大趋势。然而由于风电场出力具有显著的随机性和间歇性,风电的大量接入将对电力系统保持可靠稳定供电的能力提出挑战。近年来电力物联网技术的发展和电力市场化改革的推进,也助推了需求侧响应的相关研究和实践。以温控负荷为主要资源的需求侧响应,是提高接入风电的电力系统保持供需平衡能力的有效手段。然而温控负荷在响应特性上也具有一定不确定性。本文主要通过数据驱动算法,对需求侧的温控负荷与供电侧的风电场的不确定性进行评估,进而在考虑供需侧不确定性的前提下对电力系统进行可靠性分析,主要研究内容如下:

(1) 提出了温控负荷集群备用容量估计算法。首先基于电热力学模型实现对温控负荷个体的运行模式分析及建模,然后在考虑温控负荷参与需求侧响应时的模型参数和响应特性具有部分信息不确定性的前提下,提出基于 K-means 聚类算法和混合高斯模型的温控负荷聚合算法,实现对温控负荷集群的备用容量估计。

(2) 提出了多风电场风速概率分布估计算法。在考虑风电出力波动性的前提下,基于生成对抗网络框架,通过历史风电数据对二元人工神经网络进行博弈训练,在不依赖先验知识的条件下,实现了对风电场风速的概率分布估计以及多个邻近风电场出力的空间相关性估计。

(3) 提出了考虑温控负荷和风电不确定性的电力系统可靠性分析算法。首先基于前两个研究点对供需侧不确定性的处理,通过数据驱动方法对电力系统可靠性分析问题进行重构以解决交叉熵重要度采样算法的使用障碍,然后通过系统状态采样、系统最优负荷切除计算以及相关可靠性指标计算,实现了对考虑温控负荷和风电不确定性的电力系统可靠性的高效分析评估。

关键词: 电力系统可靠性分析; 温控负荷; 风力发电; 数据驱动;

Abstract

Due to the low carbon emission and environment concerns, the development of clean energy such as wind power has drawn great attention and become a big trend in recent years. However, the intermittent and stochastic nature of wind power can have a negative impact on the power supply reliability of power systems. Recently, with the development of the Power Internet of Things and the power system marketization reform, the research and practice of demand response have been promoted greatly. Thermostatically controlled loads (TCLs) are one of the most important resources of demand response, which can significantly mitigate the power imbalance caused by wind power fluctuations. However uncertainty also exists in the response of TCLs. This paper measures the uncertainty of TCLs and wind power by data-driven algorithms, based on which the power system reliability is also analyzed. The main research content is as follows.

(1) The reserve capacity evaluation method of aggregated TCLs is proposed. The operation mode of individual TCL is firstly analyzed and modelled based on the thermal and electrical model. Later considering the partial uncertainty of the model parameters and response characteristic of TCLs, the aggregation algorithm is proposed utilizing K-means clustering and Gaussian mixture model. Based on the aggregation algorithm the reserve capacity of aggregated TCLs is evaluated.

(2) The wind speed probability distribution estimation method in multiple wind farms is proposed. Based on the framework of generative adversarial networks, the artificial neural networks with two-player game are trained utilizing historical wind power data. The probability distribution of wind speed in multiple wind farms as well as the spatial correlation among them are estimated without any prior knowledge.

(3) The power system reliability analysis method is proposed considering the uncertainty of TCLs and wind power. Based on the uncertainty evaluation in both demand and supply sides proposed in the research above, the power system reliability

analysis problem is reconstructed by data-driven method so that the Cross-entropy importance sampling algorithm can be utilized. Later with system state sampling, optimal load curtailment calculation and related reliability indices calculation, the reliability of power systems considering the uncertainty of TCLs and wind power is analyzed and evaluated with high efficiency.

Keywords: power system reliability analysis; thermostatically controlled loads; wind generation; data-driven.

目 录

致 谢.....	I
摘 要.....	II
Abstract.....	III
插图附表清单.....	VII
1. 绪论.....	1
1.1. 课题研究的目的与意义	1
1.2. 国内外研究背景及现状	2
1.2.1. 需求侧响应与温控负荷研究现状	2
1.2.2. 风电接入评估研究现状	3
1.2.3. 电力系统可靠性评分析研究现状	4
1.2.4. 数据驱动算法研究现状	6
1.3. 本文的主要工作与章节安排	7
2. 需求侧不确定性—温控负荷集群备用容量估计.....	9
2.1. 引言	9
2.2. TCLs 运行分析与建模	10
2.3. TCLs 集群备用容量估计	11
2.3.1. 混合高斯模型	12
2.3.2. K-means 聚类.....	14
2.4. 算例分析	16
2.5. 本章小结	17
3. 供电侧不确定性—多风电场风速概率分布估计.....	19
3.1. 引言	19
3.2. 生成对抗网络模型	20
3.2.1. 对抗框架	20
3.2.2. 数学模型	21
3.3. 多风电场风速联合概率分布估计	22
3.3.1. 基于梯度法的 GANs 迭代训练算法	23

3.3.2. 相关性分析	24
3.4. 算例分析	25
3.4.1. 数据与参数概况	25
3.4.2. 风速概率分布估计结果	26
3.5. 本章小结	31
4. 考虑供需侧不确定性的电力系统可靠性分析.....	33
4.1. 引言	33
4.2. 交叉熵重要度采样算法	33
4.3. 考虑供需双侧不确定性的电力系统可靠性分析	36
4.3.1. 问题构建	37
4.3.2. 可靠性分析	39
4.4. 算例分析	43
4.4.1. 系统概述	43
4.4.2. 可靠性分析结果	44
4.5. 本章小结	46
5. 总结与展望.....	47
5.1. 对本文的总结	47
5.2. 下一步的研究展望	48
参考文献.....	49
作者简介.....	59
攻读硕士学位期间的学术成果.....	59

插图附表清单

图 1-1 人工神经网络示意图	6
图 1-2 本文研究框架示意图	7
图 2-1 TCLs 运行模型示意图	11
图 2-2 TCLs 模型参数分布示意图	16
图 2-3 TCLs 集群备用容量估计结果	17
图 3-1 GANs 模型基本框架	19
图 3-2 GANs 估计风速相关性示意图	23
图 3-3 单个风电场风速概率分布结果估计	26
图 3-4 风电场风速估计概率与真实概率之间的 KL 距离	27
图 3-5 风电场风速相关性估计结果	28
图 4-1 数据驱动的电力系统可靠性分析流程	34
图 4-2 基于 CE 算法和 GANs 处理风速隐式分布问题示意图	38
图 4-3 算例系统拓扑图	41
表 4-1 系统可靠性指标计算结果	42
图 4-4 风电场容量对系统可靠性影响示意图	42
图 4-5 TCLs 对系统可靠性影响示意图	43

1. 绪论

1.1. 课题研究的目的与意义

电力作为国民经济的重要命脉,是支撑国家经济发展和社会稳定运行的重要基础,因此保证电力系统的安全稳定运行至关重要^[1,2]。电力系统可靠性分析是对电力系统的稳定可靠供电能力的量化评估,对电网的规划和运行具有重要指导作用^[3]。近年来,随着全球气候恶化的威胁日益凸显,世界各国对能源领域的节能减排、低碳环保的要求也越来越高,大力发展风电等清洁能源已经成为能源领域发展的重要战略^[4]。提高新能源机组渗透率,降低传统火电机组占比也已经成为电力行业的发展趋势^[5]。根据全球风能委员会提供的相关新能源统计数据,截至 2018 年底,全球风电总装机容量为 591GW,与 2017 年相比增长率为 9.6%,而预计到 2022 年底总装机容量将达到 840GW^[6]。中国具有丰富的风力资源,目前风电已经成为中国的第三大电源,仅在煤电与水电之后^[7]。国家在相关的能源政策文件中也指出,为推动能源结构调整,将在“十三五”规划期间大力建设风电,鼓励各级地方政府对风力发电的资金支持和补贴并加强政策保障,计划在 2020 年实现 275GW 的全国风电总装机容量^[8]。

尽管风电能够为电力系统带来显著的减排效益,但是风电机组并网容量的增加,将为电力系统的安全可靠运行带来巨大压力。由于风电机组的运行与其所在区域的风速水平等自然环境条件紧密关联,风电机组的出力具有很大的不确定性,在不同时刻其供电功率可能相差甚远,因此大量引入风电机组大大提高了电力系统供电侧供电能力的不确定性,为电力系统保持供需平衡提出巨大挑战,提高了电力系统规划和运行的难度^[9,10]。有必要开展对电力系统中的风电场的出力波动特性的研究分析与评估,并以此为基础进一步评估其对电力系统可靠性的影响。

随着电力物联网技术的迅速发展和电力市场化改革的大力推进,需求侧响应也在电力领域获得了广泛关注^[11,12]。国家在电力体制改革的相关政策文件中指出,必须重点关注需求侧响应对确保电力系统供需平衡的重要作用^[13],这也表明了国家对需求侧响应推行的积极态度。参与需求侧响应的负荷,在电网公司的

指引下调整其自身用电行为,实现负荷侧功率消耗的优化,从而为电力系统的供需平衡提供支撑。以热泵、空调为典型代表的温控负荷,是参与需求侧响应最主要的需求侧资源,具有为电力系统提供备用的巨大潜力^[14]。因此以温控负荷为主要资源的需求侧响应,是在风电大量并网的背景下,缓解风电机组出力波动,提高电力系统可靠性的有效手段。但是由于电网公司通常无法获取系统中所有温控负荷的模型参数与响应特性,因此在评估温控负荷备用容量以及其对系统可靠性的影响时存在一定的信息不确定性。

近年来以机器学习为代表性技术的数据驱动算法的迅速发展,为电力系统复杂问题的分析,尤其是难以通过传统的模型驱动方法解决的领域,提供了新思路和新方法。本文将开展对温控负荷需求侧响应和风电的不确定性及其对电力系统可靠性影响的研究,基于数据驱动算法对温控负荷和风电场出力的随机特性进行分析、建模及评估,并以此为基础,在综合考虑两者的影响因素的前提下,提出对应的电力系统可靠性分析方法,为电力系统的规划与运行提供参考和指导。

1.2. 国内外研究背景及现状

1.2.1. 需求侧响应与温控负荷研究现状

需求侧响应是优化用电模式、提高电力系统保持供需平衡能力的重要手段^[15]。根据美国联邦能源管理委员会给出的定义,需求侧响应是指在市场尖峰电价下或者电力系统面临可靠性问题时,终端用户根据电价信号或者激励信号调整其自身用电行为^[16]。目前,有不少欧美国家已经开始实行需求侧响应,例如美国的新英格兰、纽约、加州等多个电力市场均有相关的需求侧响应计划^[17];荷兰已经通过推行基于市场电价的需求侧响应计划,成功为电力系统提供充裕的运行备用,并计划在 2020 年之前逐步扩大其需求侧响应计划的推广范围^[18]。随着中国电力市场化改革的积极推进,需求侧响应在国内的推行也引发了广泛的讨论和研究,例如文献[19]结合了国家电改政策对需求侧响应的交易模式和运营计划进行了分析;文献[20]研究了需求侧响应对电力市场的影响以及发展需求侧响应的商业模式,并对需求侧响应在中国推行的机遇与挑战进行了分析;文献[21]对需求侧响应在上海、佛山、苏州等几个试点城市的实施情况进行了调研和分析。

温控负荷作为最主要的需求侧响应资源,也吸引了较多国内外学者的关注与

研究。对于温控负荷建模，通常需要考虑其电热力学特性，目前主流的方法是基于文献[22]，[23]提出的等效热参数模型，该模型类比电路中的电导电容概念，提出等效热导参数和等效热容参数，以热流图的形式构建室内温度变化的动态过程，可通过常微分方程描述并求解温控负荷与室内外进行热量交换的模型。在等效热参数模型的基础上，文献[24]-[26]基于温控负荷的动态特性对温控负荷的控制策略进行了研究和仿真，文献[27]则在考虑市场拥塞的情况下，基于对偶分解算法提出了温控负荷参与需求侧响应的优化方案。此外，文献[28]，[29]基于马尔科夫链提出了温控负荷所在房间温度变化转移概率的计算方法，构建了温控负荷状态估计的算法。对于温控负荷参与需求侧响应的备用估计，文献[30]指出温控负荷集群中不同个体在参与需求侧响应时存在行为差异，必须纳入其备用估计的影响因素中；文献[31]对温控负荷对电价信号的响应程度进行了分析和研究；文献[32]通过仿真模型对温控负荷集群的备用容量进行计算，文献[33]则通过模糊模型实现对温控负荷集群的建模和备用容量评估。但是目前在考虑电网公司只能获取部分温控负荷个体的响应特性和模型参数的前提下，对包含温控负荷参与需求侧响应的电力系统可靠性分析的研究还比较少，有必要计及这种信息不确定性开展相关的工作。

1.2.2 风电接入评估研究现状

标题不太好，“风电出力不确定性与可靠性评估研究现状”

由于其出力的不确定性，风电大量接入电力系统后将对电力系统保持供需平衡造成巨大压力。因此国内外学者针对风电出力的不确定性及其并网对电力系统的影响开展了许多相关研究。目前对接入风电的电力系统的可靠性分析研究，大致可以分为解析法和模拟法两大类。解析法的主要思想是根据研究对象的内部机理或运行特点构建对应的数学模型，并基于该模型对研究对象的可靠性进行描述及评估计算。基于解析法，文献[34]利用通用生成函数（Universal generating function, UGF）建立了风电场及整个电力系统的可靠性模型，实现了对高风电渗透率的电力系统的可靠性分析评估及备用容量规划；文献[35]考虑风力波动和机组故障，利用停电概率和频率表方法建立了含风电机组的电力系统的可靠性模型；文献[36，37]考虑了系统的运行环境和运行条件，通过频率和持续时间法（Frequency and duration approach）建立风电并网后的电力系统的多状态模型，

这段很长。
解析法和模拟法
可以分两段。

并计算出多个状态之间的转移概率以用于可靠性和系统充裕度的计算。解析法能够建立系统运行和其它动态过程的具体数学模型，物理意义清晰明确，但是对于复杂程度高的系统，其建模的难度以及计算的复杂度也大大提高^[38]。模拟法则主要通过对系统运行状态的抽样及事件模拟实现对不确定性因素及系统可靠性的评估。模拟法可以在无需掌握系统模型的前提下实现对任意系统的特征估计，因此具有广泛的适用性，是评估复杂随机系统可靠性的主流方法^[39]。利用模拟法评估风电接入对电力系统可靠性的影响，通常的方法是对风电出力的概率分布进行估计，然后通过蒙特卡洛模拟法进行系统可靠性分析评估^[40, 41]。其中 Weibull 函数是风电概率分布中使用最为广泛的方法之一^[42-45]。此外，文献[46]基于聚类算法和蒙特卡洛模拟法对接入风电和特高压直流的电力系统可靠性进行评估，文献[47]则基于概率分布估计和模拟法提出风电机组的聚合模型，并基于可靠性指标及节点电压稳定水平提出电力系统最优风电消纳容量估计算法。

这里分段

另外，在进行概率分布估计时通常需要考虑多个随机变量之间是否存在相关性。文献[48]通过对德国、英国等多个欧盟国家的风电数据进行统计学分析，指出邻近风电场之间的风速具有显著相关性，因此在评估风电出力不确定性对电力系统可靠性的影响时，需要对不同风电场之间的相关性进行分析研究。在过往相关的工作中，研究者采用相关性系数对风速的相关性进行分析^[49,50]，但是这些方法无法获取相关的数据概率分布信息。此外，Copula 算法是目前被广泛应用于评估多风电场联合概率分布的方法。文献[51]通过基于 Copula 算法的风速相关性评估结果，提出了降低含风电的电力系统运行风险的方法；文献[52]通过 Weibull 分布和 Copula 算法建立风电场的可靠性模型；文献[53]利用 Copula 算法评估了风电接入对瑞士电网以及欧洲互联电网的影响。但是，目前对风速相关性的研究及其对电力系统可靠性的评估，往往需要依赖一定的先验知识或者无法获取显式概率分布，仍需对相关评估算法进行深入研究。

1.2.3. 电力系统可靠性评分分析研究现状

这里少了个字，“可靠性评估？分析？不能是评分”

电力系统可靠性分析评估的概念最早是由加拿大的 Roy Billinton 教授提出的，它是对电力系统保持可靠供电能力的分析评估，通常利用基于概率的指标进行表示^[54]。经过多年的发展，国内外学者已经在电力系统可靠性评分分析估领域

积累了许多研究成果^[55-59]。对于计及需求侧响应的接入风电的电力系统,文献[60]分析了风电机组和需求侧响应对电力系统频率的影响作用,提出了需求侧响应、风电、火电机组联合参与系统调频的控制方法,文献[61]基于弹性矩阵理论分析了参与需求侧响应的负荷特性,利用分层优化模型及算法评估了需求侧响应对风电出力波动的效应,文献[62]分析了市场电价对电动汽车等机会型负荷的影响,并在不同时间尺度下分析了机会型负荷对含风电的电力系统运行计划制定的影响。然而目前在考虑温控负荷模型参数及风电出力不确定性的前提下,对电力系统进行可靠性分析评估的研究仍然比较少。

在电力系统可靠性分析评估算法方面,如前文所述,以蒙特卡洛模拟法为代表的模拟法是分析评估复杂系统可靠性问题的主要方法。但是蒙特卡洛模拟法的一个缺点是其算法需要较大规模的采样数量以获得收敛的分析评估结果,尤其是在评估低概率事件的情况下。文献[63]指出,当电力系统具有相对较高的可靠性时,采用蒙特卡洛模拟法会导致较高的计算负担。为了提高模拟法的计算效率,一些学者进行了相关的研究工作^[64-66]。文献[64]采用了状态空间分解算法降低了可靠性分析评估问题的维度以提高蒙特卡洛模拟法的计算效率;文献[65]提出了构建分布式并行计算的评估框架以实现模拟计算加速的方案;文献[66]采用了交叉熵重要度采样算法并指出这是提高模拟法计算效率的有效方法。近年来已有不少电力系统相关研究工作采用了交叉熵重要度采样算法^[67-69]。文献[67]通过交叉熵重要度采样算法对电力系统的旋转备用要求进行评估,文献[68]提出了基于交叉熵重要度采样算法的三阶段采样算法以实现电力系统的短期可靠性的分析评估,文献[69]指出在复杂电力系统可靠性分析评估中交叉熵重要度采样算法的计算性能优于蒙特卡洛模拟法。尽管交叉熵重要度采样算法具有较优的计算性能,但是当使用该算法处理具有隐式概率分布的数据问题时会遭遇障碍。文献[66]指出为了顺利采用交叉熵重要度采样算法,其模拟的问题的概率分布必须属于已知的参数化分布族。为了解决上述障碍,文献[69]采用了高斯分布对电力系统中的负荷进行概率估计,文献[70]则在其电力系统充裕度评估算法中引入了 Weibull 分布和对数正态分布以保证交叉熵重要度采样算法的顺利使用。然而这些算法都在数据分布问题上引入了比较强的先验知识(例如对数据分布类型的假定),这

此处最好分段

在某些情况下可能引入较大误差，影响了算法的通用性^[71]。

1.2.4. 数据驱动算法研究现状

数据驱动是与传统的模型驱动相对应的概念。数据驱动算法从基于测量或者仿真获取的数据中提取挖掘被研究对象的特征，从而建立其算法的知识体系及问题的求解，这与基于物理数学建模的模型驱动算法具有截然不同的思路^[72]。近年来，随着机器学习等领域深入研究和迅速发展，数据驱动算法被认为在解决传统建模难以进行的复杂问题上具有巨大的潜力。机器学习以统计学为基础，通过对数据的学习建立用于解决特定任务的模型，通常可以分为监督学习、无监督学习及增强学习^[73]。除了较早提出的 Logistic 回归、支持向量机（Support vector machine, SVM）、朴素贝叶斯法等经典算法^[73,74]，随着计算资源的发展和相关理论研究的深入，以多层人工神经网络（Artificial Neural Network, ANN）为基础的深度学习也日益获得了众多研究者的重点关注^[75]。图 1-1 展示了一个典型的 ANN 模型，它由输入层、隐藏层和输出层相连构成，每一层则都由数个神经元组成，每一个神经元以其前一层神经元输出为输入，根据网络参数及激活函数确定其输出并传递到下一层神经元。ANN 实质上是一个非线性函数，实现了从输入到输出的映射，理论上当 ANN 具有足够大的深度时，通过数据训练确定其网络参数后可以逼近任何函数，具有很大的特征学习能力，因此在数据驱动算法中获得了广泛应用。

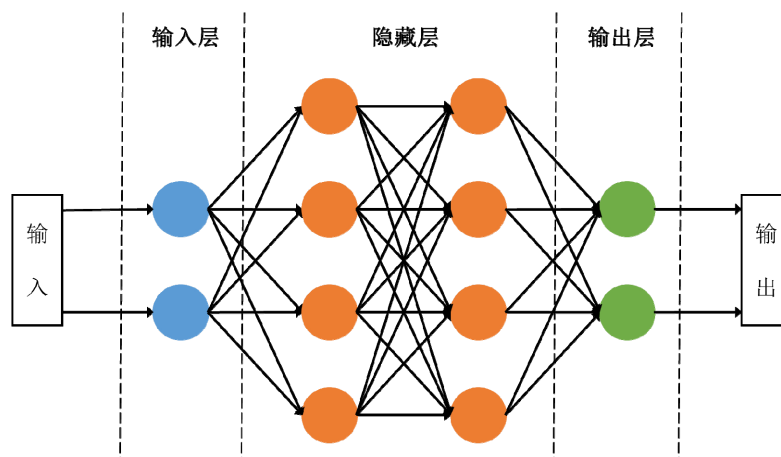


图 1-1 人工神经网络示意图

目前已有较多学者对数据驱动算法在电力系统中的应用进行研究，分析利用

数据驱动方法解决电力系统中的复杂模型或者复杂特征问题的潜力^[76-80]。文献[76]利用堆栈自动编码器（Stacked auto-encoder）和 ANN 实现对电力系统的故障诊断，其最终搭建的数据驱动诊断模型能够基于系统状态计算发生故障的种类及概率；文献[77]利用机器学习中的自然语言处理技术，提出了基于语义分析的电力系统缺陷文本挖掘算法，实现了电力系统缺陷自动分类；文献[78]基于极限学习机算法，结合电力系统频率响应模型，提出了电力系统频率稳定控制算法；文献[79]基于数据驱动算法对商业楼宇用电负荷进行研究，通过关键特征提取和 ANN 实现了用电负荷预测；文献[80]结合专家系统与强化学习算法，提出了电力市场环境下的温控负荷最优控制策略。可以看出数据驱动算法在电力系统领域具有广阔的应用空间，在未来的研究工作中具有深入探索的价值。

1.3. 本文的主要工作与章节安排

风电的大量接入以及温控负荷需求侧响应的引入，对电力系统的可靠性具有不可忽视的影响。本文通过数据驱动的方法，对需求侧的温控负荷不确定性和供电侧的风电场出力不确定性进行建模与评估，并以此为基础进行电力系统可靠性分析。本文的研究框架如图 1-2 所示。

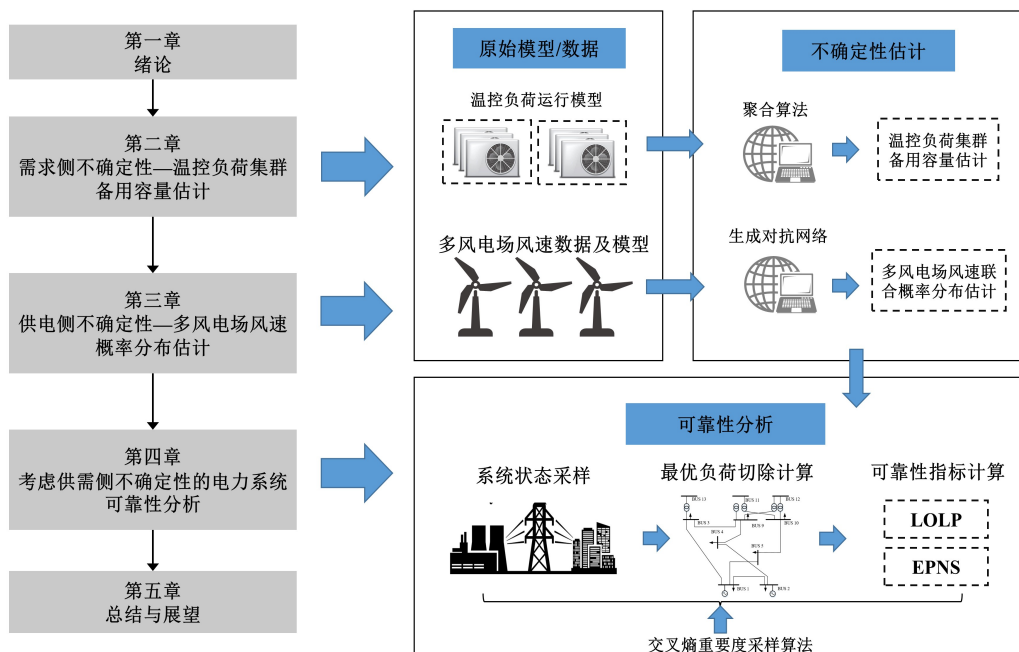


图 1-2 本文研究框架示意图

本文共分为五章，每一章的内容概述如下：

第一章为全文绪论，阐述了本文研究课题的背景、目的及意义，并对需求侧响应与温控负荷、风电接入评估、电力系统可靠性分析、数据驱动算法这几个方面的国内外研究现状进行总结与概述。

第二章对需求侧的温控负荷不确定性进行评估。本文考虑温控负荷参与需求侧响应时的模型参数和响应程度具有部分信息不确定性，首先根据电热力学方程建立温控负荷个体的运行模型，然后提出基于 K-means 聚类算法和混合高斯模型的数据驱动温控负荷聚合算法，实现对温控负荷集群的备用容量估计，为后续电力系统可靠性分析提供需求侧的模型基础。

第三章对供电侧的风电出力不确定性进行评估。本文考虑风电出力的波动性，提出了基于生成对抗网络的多风电场风速概率分布估计方法，在无需引入先验知识的条件下，通过历史风电数据开展基于对抗的 ANN 训练，实现对风电场风速概率分布以及邻近风电场空间相关性的估计，以此作为后续电力系统可靠性分析的基础。

个人疑问：无需引入先验知识，“通过历史风电数据开展XXX”这里的历史数据不算先验知识？

第四章对考虑温控负荷和风电不确定性的电力系统进行可靠性分析。本章基于前两章对供需双侧不确定性的处理，通过数据驱动方法将电力系统可靠性分析问题构建为可以直接采用交叉熵重要度采样算法进行处理的分析评估问题，并通过系统状态采样、最优负荷切除计算以及可靠性指标计算，实现对电力系统可靠性的高效分析评估。

第五章对全文的研究内容进行了总结，并对本文所提的研究方法及后续研究方向进行了展望。

2. 需求侧不确定性—温控负荷集群备用容量估计

2.1. 引言

近年来,随着电力系统运行的复杂性和随机性增大(例如新能源机组并网容量的提高)以及电力市场化的逐步推进,需求侧响应引起了学界和工业界的重点关注^[11,12]。在电力市场环境中,电力供应商通过调整实时电价并向加入需求侧响应计划的用户发送电价信号,引导用户调整自身用电行为,从而调整电力系统需求侧的整体负荷消耗功率,这是一种常见的需求侧响应模式^[17]。这些能够在需求侧响应中调整自身用电功率的负荷,称为需求侧响应资源。

温控负荷(Thermostatically controlled loads, TCLs)是电力系统中最主要的需求侧响应资源之一,被认为具有巨大的提供电力系统运行备用的潜力,能够提高电力系统应对各种复杂运行状况的能力^[32]。TCLs泛指通过消耗电功率控制室内温度的负荷,通常有制冷或者制热两种类型,空调即为典型的TCLs。TCLs作为需求侧响应资源,具有3个主要优势:(1)TCLs的消耗功率在电力系统负荷中具有很大的占比,具有巨大的可调度功率容量^[81](2)TCLs与传统发电机组相比,具有更快的响应速度。文献^[82]指出TCLs的响应时间常数可以达到传统火电发电机组的1/10。(3)由于室内温度具有惯性,短时间内对TCLs用电功率的调整,不会对用户的使用体验造成太大的负面影响^[83]。

TCLs参与需求侧响应,能够在供电紧张时提供运行备用,从而提高电力系统的可靠性。因此本文在进行电力系统可靠性分析时,将在需求侧考虑TCLs的响应作用,并通过估计TCLs集群的备用容量来评估其对系统可靠性的影响。在一个电力系统的需求侧响应计划中一个TCLs集群一般是由数量庞大的TCLs个体组成的,且不同TCLs个体之间可能具有不同的模型参数和响应特性。但是通常电网公司只能采集掌握到部分TCLs个体的模型参数和响应特性,要完全掌握巨大数量的TCLs个体的响应相关信息往往是不现实的。因此在进行TCLs备用容量评估和系统可靠性分析时存在一定的信息不确定性。为了解决上述问题,本文首先基于电热模型建立TCLs个体的运行模型,然后提出基于K-means聚类算法的TCLs聚合算法,通过部分TCLs个体的响应特性和模型参数实现对TCLs

集群的备用容量估计，为后续电力系统可靠性分析提供需求侧的模型基础，并通过算例分析验证所提方法的有效性。

2.2. TCLs运行分析与建模

本节将对 TCLs 的运行模式进行分析并基于电热模型对其运行功率进行计算，为后续 TCLs 聚合算法提供基础。假定 TCLs 工作在制冷模式。根据热力学定律，TCLs 所在的房间温度可以用下式表示^[32]：

$$T_r(t) - T_r(t + \Delta t) = \int_t^{t+\Delta t} \frac{P_H(t) - P_{ex}(t)}{CM_r} dt \quad (1.1)$$

其中 $T_r(t)$ 表示 t 时刻房间的温度， CM_r 表示房间的热质量 (Thermal mass)， $P_H(t)$ 表示在 t 时刻 TCLs 的制冷热功率， $P_{ex}(t)$ 表示在 t 时刻房间与室外的交换功率， Δt 表示一段时间间隔。

$P_H(t)$ 和 $P_{ex}(t)$ 分别可以由以下公式表示：

$$P_H(t) = \eta_r \cdot P_T(t) \quad (1.2)$$

$$P_{ex}(t) = G_r(T_{out} - T_r(t)) \quad (1.3)$$

其中 $P_T(t)$ 表示在 t 时刻 TCLs 的运行功率， η_r 表示 TCLs 的能效比 (Energy efficiency ratio)， G_r 表示室内外空气之间的等效热导， T_{out} 表示室外温度。

TCLs 的运行模式如图 2-1 所示^[83]。图 2-1 的上半部分表示室内温度的变化过程，其中 T_{set} 和 T_d 分别表示 TCLs 的设定温度和温度死区，下半部分表示 TCLs 运行功率的变化过程，其中 P_{T0} 表示额定功率。TCLs 的运行模式是一个循环过程：当室内温度低于 $T_{set} - T_d$ 时，TCLs 停止运行，运行功率为 0，因此室内温度逐渐上升；当室内温度高于 $T_{set} + T_d$ 时，TCLs 以 P_{T0} 开始运行，此时室内温度逐渐下降，当室内温度达到 $T_{set} - T_d$ 时结束一个循环。因此 TCLs 的运行功率 $P_T(t)$ 可用下式表达：

$$P_T(t) = \begin{cases} P_{T0}, & T_r(t) \geq T_{set} + T_d \\ 0, & T_r(t) \leq T_{set} - T_d \\ \text{保持不变, 其它} \end{cases} \quad (1.4)$$

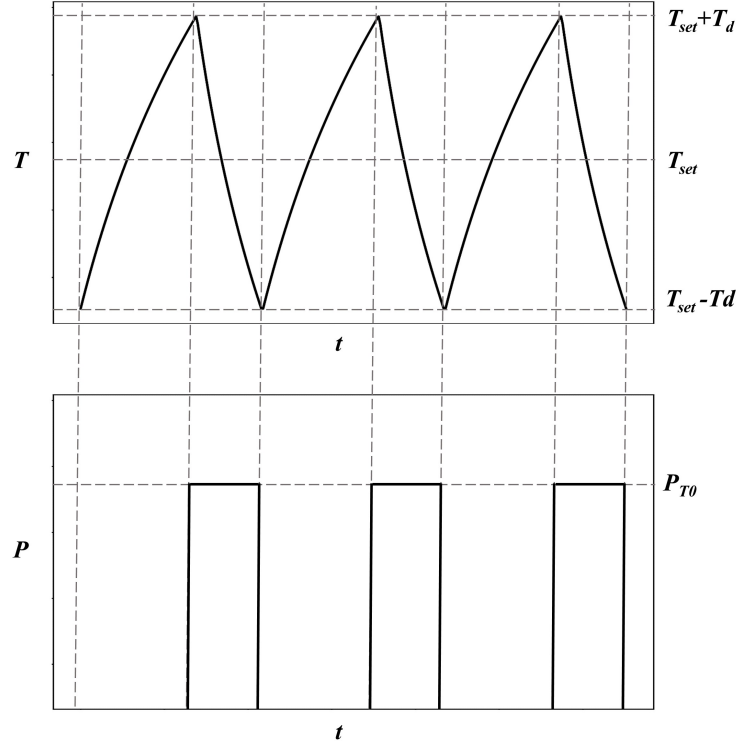


图 2-1 TCLs 运行模型示意图

由式(1.1)—(1.4)可以推出, TCLs 的平均运行功率为:

$$P_{Ta} = \frac{G_r(T_{out} - T_{set})}{\eta_r} \quad (1.5)$$

2.3. TCLs 集群备用容量估计

在本文中, 需求侧的不确定性主要是由不同 TCLs 个体的模型参数(包括电热力学参数及对电价的响应程度)的不同导致的。本节将针对该不确定性进行建模, 通过数据驱动的聚合算法对参与需求侧响应的 TCLs 集群备用容量进行估计。

在参与需求侧响应的过程中, TCLs 根据电价控制信号调整设定温度, 进而改变其运行功率, 从而实现系统负荷用电功率的调整。 T_{set} 通常与电价水平相关, 可通过下式进行计算^[31,33]:

$$T_{set} = \left\lfloor T_d + \frac{T_{\max} - T_d}{1 + e^{-\alpha p}} \right\rfloor \quad (1.6)$$

$$T_d = 2T_{\min} - T_{\max}$$

其中 $T_{\min}$ 和 $T_{\max}$ 分别表示 TCLs 的最低和最高设定温度, p 表示电价; α 表示 TCLs 对电价的响应程度, $\lfloor \cdot \rfloor$ 表示向下取整。

由于单一 TCLs 的功率相对电力系统而言非常小, 因此只有多个个体构成的 TCLs 集群才能充分发挥需求侧响应的作用。由式(1.5)和(1.6)可以看出, 单一 TCLs 的平均运行功率与其自身电热力学参数以及对电价的响应程度有关。然而不同 TCLs 个体之间存在差异, 并且集群温控负荷中个体数目庞大, 难以通过逐一测量计算每个个体的参数来获取集群温控负荷的备用容量。本文假定不同 TCLs 个体可能具有不同的 G_r 和 α 。对于一个个体总数为 N_r 的 TCLs 集群, 假定其中有 N_s 个个体的模型参数已知, 并用 $\Phi_T = \{(G_r^1, \alpha^1), (G_r^2, \alpha^2), \dots, (G_r^{N_s}, \alpha^{N_s})\}$ 进行表示, 其中 G_r^i 和 α^i 分别为第 i 个 TCLs 个体的对应模型参数。本节基于这部分已知的模型参数, 通过混合高斯模型和 K-means 聚类这两种数据驱动的聚合算法方案对集群温控负荷备用容量进行评估, 实现对需求侧的不确定性建模。

其实对个体的影响因素很多, 本文选择了这样2个主要因素, 进行聚类分析, 简化问题。否则维度太高, 聚类也会复杂。

2.3.1. 混合高斯模型

混合高斯模型是一种通过多个高斯分布函数实现概率分布估计的算法^[84]。本节首先通过混合高斯模型对 G_r 和 α 的概率分布进行估计, 然后基于数学期望计算 TCLs 群体的备用容量。

混合高斯模型假定 G_r 的概率分布为:

$$f_{G_r}(x; \mathbf{w}_G, \boldsymbol{\mu}_G, \boldsymbol{\sigma}_G) = \sum_{i=1}^{k_{Gr}} w_{Gi} \phi(x; \mu_{Gi}, \sigma_{Gi}) \quad (1.7)$$

$$\sum_{i=1}^{k_{Gr}} w_{Gi} = 1$$

其中 $\phi(x; \mu_{Gi}, \sigma_{Gi})$ 表示均值为 μ_{Gi} , 标准差为 σ_{Gi} 的高斯分布函数, w_{Gi} 表示 $\phi(x; \mu_{Gi}, \sigma_{Gi})$ 对应的权重系数, $\mathbf{w}_G = \{w_{G1}, w_{G2}, \dots, w_{Gk_{Gr}}\}$, $\boldsymbol{\mu}_G = \{\mu_{G1}, \mu_{G2}, \dots, \mu_{Gk_{Gr}}\}$, $\boldsymbol{\sigma}_G = \{\sigma_{G1}, \sigma_{G2}, \dots, \sigma_{Gk_{Gr}}\}$, k_{Gr} 表示高斯分布函数总数。由式(1.7)可以看出, 混合

高斯模型首先基于权重系数计算 $\mathbf{x}$ 属于每个高斯分布的概率，然后再通过高斯分布概率加权求和确定 $\mathbf{x}$ 的概率。因此混合高斯模型需要通过计算确定的参数为 $\mathbf{w}_G$ ， $\boldsymbol{\mu}_G$ 和 $\boldsymbol{\sigma}_G$ ，可以通过极大似然估计（Maximum likelihood estimation, MLE）实现^[73]。MLE 的核心思想是计算出一组 $\mathbf{w}_G$ ， $\boldsymbol{\mu}_G$ 和 $\boldsymbol{\sigma}_G$ 的值，使得在通过这组参数确定的概率分布函数下，已知数据出现的概率最大：

$$\begin{aligned} (\mathbf{w}_G, \boldsymbol{\mu}_G, \boldsymbol{\sigma}_G) &= \arg \max_{(\mathbf{w}_G, \boldsymbol{\mu}_G, \boldsymbol{\sigma}_G)} \prod_{i=1}^{N_s} f_{G_r}(G_r^i; \mathbf{w}_G, \boldsymbol{\mu}_G, \boldsymbol{\sigma}_G) \\ &= \arg \max_{(\mathbf{w}_G, \boldsymbol{\mu}_G, \boldsymbol{\sigma}_G)} \sum_{i=1}^{N_s} \log f_{G_r}(G_r^i; \mathbf{w}_G, \boldsymbol{\mu}_G, \boldsymbol{\sigma}_G) \end{aligned} \quad (1.8)$$

由于直接求解式 (1.8) 难度较高，通常采用期望最大化（Expectation-Maximization, EM）算法^[84]进行混合高斯模型的参数求解。EM 算法是一种迭代算法，它的每一轮迭代都由期望和最大化两个步骤组成，其具体计算过程如下所示：

步骤 1：算法初始化，随机生成参数的初值 $\mathbf{w}_G^0$ ， $\boldsymbol{\mu}_G^0$ 和 $\boldsymbol{\sigma}_G^0$ 。

步骤 2：期望步骤。基于第 $t-1$ 轮迭代得到的参数结果 $\mathbf{w}_G^{t-1}$ ， $\boldsymbol{\mu}_G^{t-1}$ 和 $\boldsymbol{\sigma}_G^{t-1}$ ，可以通过贝叶斯公式计算每一个 G_r 样本属于每一个高斯分布函数的概率：

$$p_{ij} = p(i|G_r^j) = \frac{w_{Gi}^{t-1} \phi(G_r^j; \boldsymbol{\mu}_{Gi}^{t-1}, \boldsymbol{\sigma}_{Gi}^{t-1})}{\sum_{h=1}^{k_{G_r}} w_{Gh}^{t-1} \phi(G_r^j; \boldsymbol{\mu}_{Gh}^{t-1}, \boldsymbol{\sigma}_{Gh}^{t-1})} \quad (1.9)$$

其中 p_{ij} 表示 G_r^j 属于第 i 个高斯分布函数的概率。

步骤 3：最大化步骤。基于期望步骤中计算的 p_{ij} ，对本轮迭代（第 t 轮）混合高斯模型的参数进行更新：

$$w_{Gi}^t = \frac{\sum_{j=1}^{N_s} p_{ij}}{N_s} \quad (1.10)$$

$$\mu_{Gi}^t = \frac{1}{\sum_{j=1}^{N_s} p_{ij}} \sum_{j=1}^{N_s} p_{ij} G_j^r \quad (1.11)$$

$$\sigma_{Gi}^t = \sqrt{\frac{1}{\sum_{j=1}^{N_s} p_{ij}} \sum_{j=1}^{N_s} p_{ij} (G_j^r - \mu_{Gi}^t)^2} \quad (1.12)$$

其中 w_{Gi}^t , μ_{Gi}^t 和 σ_{Gi}^t 分别为 w_G^t , μ_G^t 和 σ_G^t 中的第 i 个元素。

步骤 4: 如果第 t 轮的参数计算结果与第 $t-1$ 轮的结果之差小于预设的误差范围, 则认为迭代达到收敛, 算法停止, 输出 w_G^t , μ_G^t 和 σ_G^t 作为混合高斯模型的最终参数, 否则令 $t \leftarrow t+1$ 并重复步骤 2 进入下一轮迭代。

TCLs 的 α 参数同样可以通过上述 EM 算法进行计算, 并最终确定其对应参数 w_α , μ_α 和 σ_α 。随后可以基于数学期望计算 TCLs 集群的平均运行功率, 如下式所示:

$$s \overline{P_T^M}(p) = \int \int_{G_r} f_{G_r}(x; w_G, \mu_G, \sigma_G) f_\alpha(y; w_\alpha, \mu_\alpha, \sigma_\alpha) P_{Ta} dy dx \quad (1.13)$$

其中 $\overline{P_T^M}(p)$ 表示在电价 p 下的 TCLs 集群的平均运行功率。因此, TCLs 集群基于电价控制信号能够提供的备用容量为:

$$R_T^M = N_T [\overline{P_T^M}(p_{old}) - \overline{P_T^M}(p_{new})] \quad (1.14)$$

其中 R_T 为 TCLs 集群备用容量, p_{old} 和 p_{new} 分别为电价控制信号发出前后的电价水平。

2.3.2. K-means 聚类

K-means 聚类是一种基于 $L2$ 距离的常用聚类算法^[85]。本节通过 K-means 聚类获取 Φ_T 数据集中的聚类中心, 再基于聚类中心实现对 TCLs 集群的备用容量评估。

K-means 聚类算法的核心是基于 $L2$ 距离最小化计算出最优的聚类中心。在 K-means 聚类中, Φ_T 中的每一个元素都唯一属于一个聚类中心。假设聚类中心

的总数为 N_c ，针对给定的聚类中心集合 $C = \{(\hat{G}_r^1, \hat{\alpha}^1), (\hat{G}_r^2, \hat{\alpha}^2), \dots, (\hat{G}_r^{N_c}, \hat{\alpha}^{N_c})\}$ ，定义 (G_r^i, α^i) 所属的聚类中心 $(\hat{G}_r^{(i)}, \hat{\alpha}^{(i)})$ 是 C 中距离其最近的元素，即：

$$(G_r^{(i)}, \alpha^{(i)}) = \arg \min_{(G_r, \alpha) \in C} [(G_r^i - G_r)^2 + (\alpha^i - \alpha)^2] \quad (1.15)$$

则最优聚类中心集合 C^* 可通过下式计算：

$$C^* = \arg \min_C \sum_{i=1}^{N_s} (G_r^i - G_r^{(i)})^2 + (\alpha^i - \alpha^{(i)})^2 \quad (1.16)$$

通常可以利用迭代算法计算式(1.16)的解，在每一轮迭代中，基于上一轮迭代计算的聚类中心结果确定 Φ_T 中每一个元素所属的聚类中心，然后再对聚类中心进行更新，具体如下所示：

步骤 1：算法初始化，随机生成聚类中心集合的初始值，用 C^0 表示。

步骤 2：假设第 $t-1$ 轮迭代得到的聚类中心集合为 C^{t-1} ，基于式(1.15)计算 Φ_T 中每一个元素所属的聚类中心，并用 $(\hat{G}_r^{(i)t-1}, \hat{\alpha}^{(i)t-1})$ 表示在第 $t-1$ 轮迭代中 Φ_T 的第 i 个元素所属的聚类中心。

步骤 3：更新聚类中心。通过以下公式计算本轮迭代（第 t 轮）的聚类中心集合 C^t ：

$$\hat{G}_r^{it} = \frac{\sum_{j=1}^{N_s} k_{ij} G_r^j}{\sum_{j=1}^{N_s} k_{ij}} \quad (1.17)$$

$$\hat{\alpha}^{it} = \frac{\sum_{j=1}^{N_s} k_j \alpha^j}{\sum_{j=1}^{N_s} k_j} \quad (1.18)$$

其中 $(\hat{G}_r^{it}, \hat{\alpha}^{it})$ 表示 C^t 中的第 i 个元素， k_{ij} 表示聚类中心归属系数，如果 (G_r^j, α^j) 属于聚类中心 $(\hat{G}_r^{it}, \hat{\alpha}^{it})$ ，则 k_{ij} 取值为 1，否则取 0。

步骤 4：如果第 t 轮的聚类中心计算结果与第 $t-1$ 轮的结果之差小于预设的误差范围，则认为迭代达到收敛，算法停止，输出 C^t 作为 K-means 聚类的聚类

中心集合计算结果，否则令 $t \leftarrow t + 1$ 并重复步骤 2 进入下一轮迭代。

基于 K-means 聚类最终确定的最优聚类中心 $C^* = \{(\hat{G}_r^{1*}, \hat{\alpha}^{1*}), (\hat{G}_r^{2*}, \hat{\alpha}^{2*}), \dots, (\hat{G}_r^{N_c^*}, \hat{\alpha}^{N_c^*})\}$ ，可以通过下式对 TCLs 集群的平均运行功率 $\overline{P_T^K}(p)$ 进行估计：

$$\overline{P_T^K}(p) = \frac{1}{N_s} \sum_{i=1}^{N_c} N_c^i P_{T_a}(\hat{G}_r^{i*}, \hat{\alpha}^{i*}, p) \quad (1.19)$$

其中 N_c^i 表示 Φ_T 中属于聚类中心 $(\hat{G}_r^{i*}, \hat{\alpha}^{i*})$ 的样本数量， $P_{T_a}(\hat{G}_r^{i*}, \hat{\alpha}^{i*}, p)$ 表示基于式 (1.5)，在电价 p 下通过 $\hat{G}_r^{i*}$ ， $\hat{\alpha}^{i*}$ 确定的 TCLs 个体平均运行功率，则 TCLs 集群的备用容量可以通过下式进行估计：

$$R_T^K = N_T [\overline{P_T^K}(p_{old}) - \overline{P_T^K}(p_{new})] \quad (1.20)$$

2.4. 算例分析

本节构建一个个体总数 N_T 为 30000 的 TCLs 集群，并利用所提方法对其备用容量进行估计。利用拉普拉斯分布随机生成 TCLs 集群中每一个个体的 G_r 和 α ，其中拉普拉斯分布的概率分布函数如下所示：

$$f_L(x; \delta, \tau) = \frac{1}{2\tau} e^{(-|x-\delta|/\tau)} \quad (1.21)$$

假设 G_r 有 50% 的比例满足概率分布 $f_L(x; 2.8, 0.7)$ ，另外 50% 满足概率分布 $f_L(x; 4.5, 0.7)$ （单位为 kW/°C）。类似的，假设 α 有 50% 的比例满足概率分布 $f_L(x; 3.33, 0.5)$ ，另外 50% 满足概率分布 $f_L(x; 6.66, 0.5)$ （单位为 ¥⁻¹）。 G_r 和 α 的概率分布如下图所示：

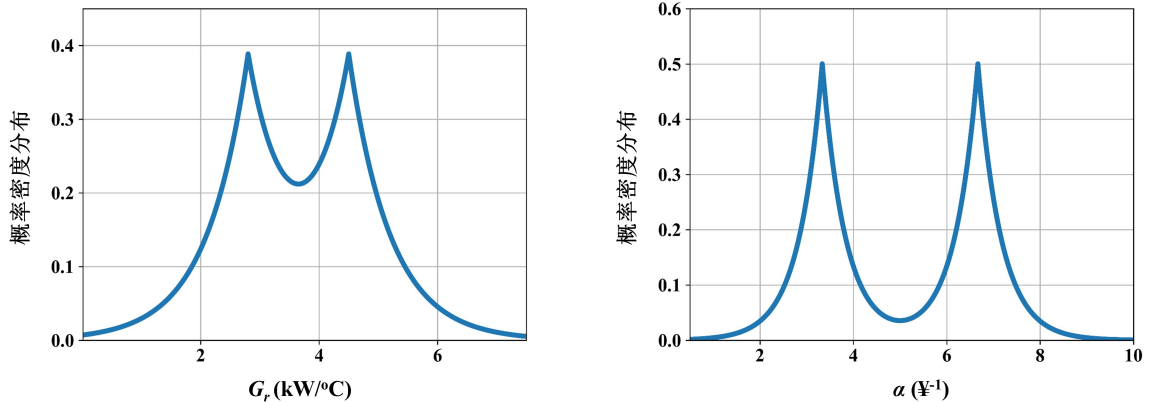


图 2-2 TCLs 模型参数分布示意图

和图片同一页

假设 TCLs 集群中模型参数已知的个体数量 N_s 为 600，室外温度为 32°C ，初始电价水平为 0.5¥ ，分别采用混合高斯模型和 K-means 聚类对 TCLs 集群的备用容量进行估计，在不同的电价控制信号下两种估计方法的计算结果如图 2-3 所示。图 2-3 的左半部分为实际备用容量以及对应的估计结果，右半部分为两种估计方法的结果与实际数据之间的误差。可以看出，两种估计方法都可以实现对 TCLs 集群的备用容量及其随电价变化趋势的大致估计，但是在不同的电价水平下 K-means 聚类的估计结果误差均小于混合高斯模型，故 K-means 聚类相比混合高斯模型具有更高的估计准确性，能够对已知数据集实现更加有效的特征提取和挖掘。

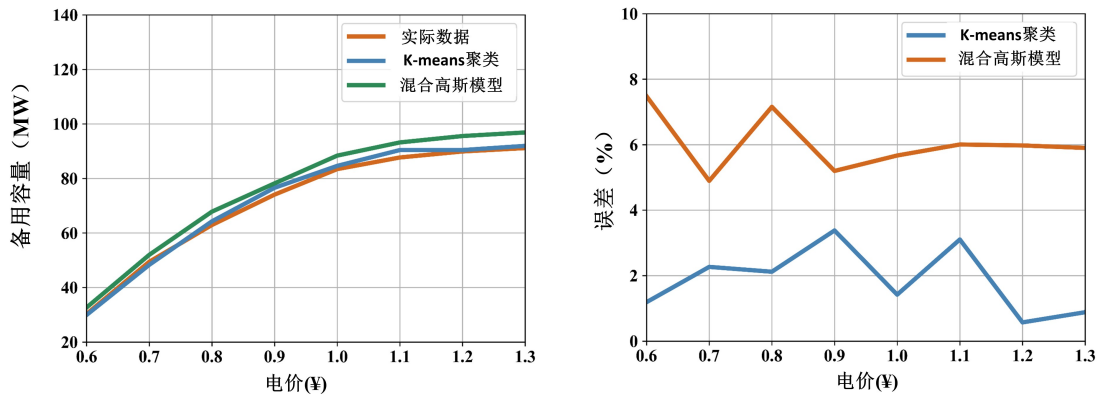


图 2-3 TCLs 集群备用容量估计结果

2.5. 本章小结

TCLs 集群在参与需求侧响应时，由于不同的 TCLs 个体具有不同的响应特性和模型参数，因此具有不同的响应程度。由于电网企业通常无法获取 TCLs 集群的所有响应相关信息和参数，因此在评估 TCLs 参与需求侧响应对电力系统可靠性的影响时，具有信息不确定性。本章首先对 TCLs 个体的运行模式和功率模型进行分析和构建，然后基于 TCLs 集群中部分已知的模型参数数据，通过混合高斯模型和 K-means 聚类这两种数据驱动的聚合算法进行 TCLs 集群备用容量估计。

算例分析表明，基于K-means聚类的聚合算法可以实现对TCLs集群备用容量的更加精确有效的估计。

3. 供电侧不确定性—多风电场风速概率分布估计

3.1. 引言

随着全球对环境保护的重视程度越来越高,风电等清洁能源在电力系统中的投入比重也日益增长^[5]。然而由于风电机组出力与所处地理区域的风速直接关联,具有很大的间歇性和随机性,大量引入风电场将对电力系统保持供需平衡提出巨大挑战,从而对电力系统可靠性造成显著影响^[9]。为了评估风电出力不确定性对电力系统可靠性的影响,目前常见的方法是首先对风电场风速的概率分布进行估计,并以此作为后续电力系统可靠性分析的基础。

一个电力系统中往往存在多个风电场同时接入,这些风电场由于所处地理位置有所不同,其风速也不完全相同,因此对风电场风速进行概率分布估计,主要有两个问题需要重点关注:

- (1) 对于单个风电场风速的概率分布估计,目前常用的方法主要为基于Weibull函数的估计等方法^[42-45];
- (2) 对多个风电场风速之间的相关性估计及联合概率分布估计。由于同一区域中的多个风电场风速之间往往存在相关性,因此不同风电场风速之间不互相独立^[48],为了保证对可靠性的准确评估,因此在进行多风电场风速分布估计的时候,必须对其联合分布进行构建,而不仅仅关注各自的边缘分布。目前常用的方法为基于Copula算法估计^[51-53]。

对于上述两个问题,目前常用的概率估计方法存在的不足之处在于它们引入了先验知识。这里所谓的先验知识,是指这些方法在进行概率估计之前,对其处理的数据事先做出了服从某些特定分布的假设。然而实际上我们无法保证数据是否真正服从预设的概率分布。因此这些先验知识的引入,导致了该估计方法在普适性上存在一定不足,无法灵活地应对不同场景下数据分布特点的差异性,挖掘数据的复杂特征。为了应对上述的不足,本文提出了基于生成对抗网络(Generative adversarial networks, GANs)^[86]的多风电场风速概率分布估计方法,在不依赖任何先验知识的前提下,通过真实历史风电数据,实现对多风电场风速联合概率分布的估计,以此作为后续电力系统可靠性分析的基础,并通过算例分

ok,回答了前面的问题

析验证该方法的有效性。

3.2. 生成对抗网络模型

3.2.1. 对抗框架

GANs 作为一种人工智能领域新兴的生成模型，其主要功能是对给定训练数据集进行概率分布估计，且可以应用于对高维度数据的联合概率分布估计。GANs 是由生成网络（Generative Network）和判别网络（Discriminative Network）构成的，其核心思想是两个不同网络在学习过程中的博弈对抗，其基本框架如图 3-1 所示^[86]。

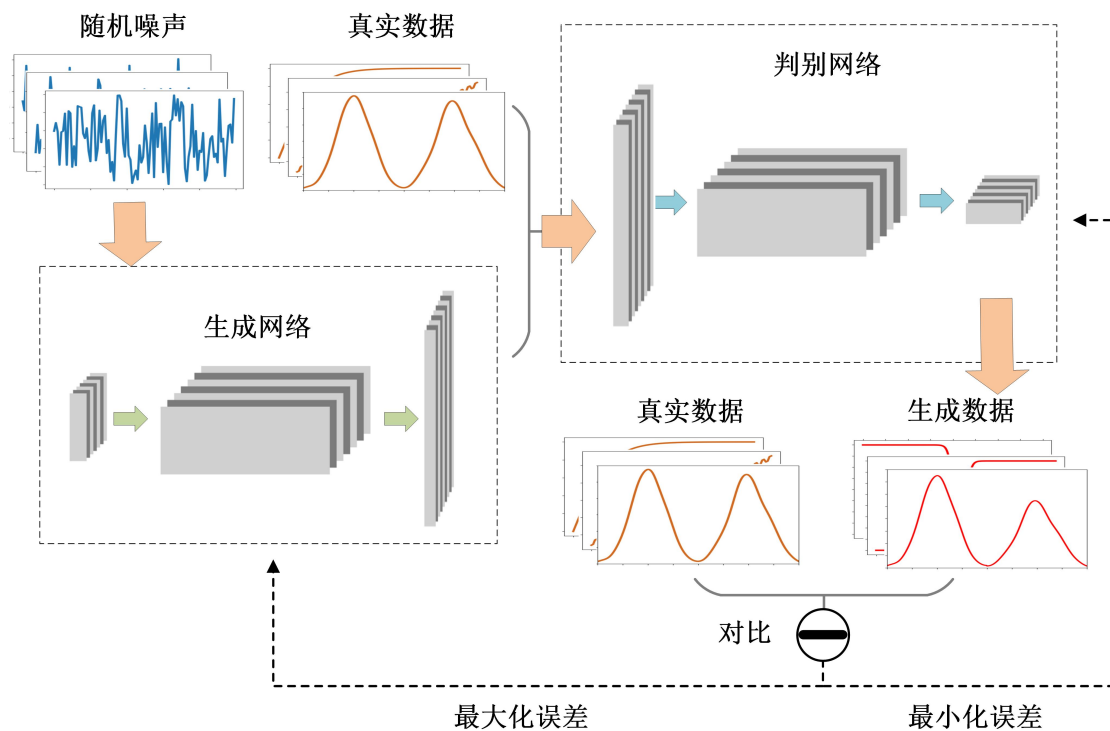


图 3-1 GANs 模型基本框架

理论上，生成网络和对抗网络可以采用多种形式的模型，但目前学界通常采用人工神经网络模型^[75]。GANs 的任务是通过生成网络逼近给定的原始数据集的概率分布。具体地，生成网络以服从某一分布的随机噪声作为输入，其任务是生成并输出逼近原始数据集分布的数据（随机噪声的分布可以任意给定）。判别网络则以原始数据集和生成网络生成的数据作为输入，其任务是一个分类问题，即判断输入的数据是来自于实际数据或是生成数据。在 GANs 的训练过程中，生成网络的训练目标是最大化判别网络的分类错误率，而判别网络的训练目标是最小

化其分类错误率。因此，生成网络和判别网络之间存在对抗的关系：即生成网络的学习方向是生成更加逼近原始数据集的分布的数据以使判别网络无法分辨出它们的差别，而判别网络的学习方向则是不把生成网络生成的数据和原始数据集混淆。通过这种对抗，生成网络和判别网络不断提升完成各自任务的性能。经过多轮训练的迭代后，生成网络将实现对原始数据集概率分布的逼近。

3.2.2. 数学模型

GANs 的训练过程，实质上是由生成网络和判别网络构成的二元极小极大博弈（Two-player minimax game），这是一个优化问题^[86]。具体地，这个优化问题的目标函数为：

$$\min_{\theta_G} \max_{\theta_D} V(D, G) = \mathbb{E}_{\mathbf{x} \sim p_{real}} [\log D(\mathbf{x}; \theta_D)] + \mathbb{E}_{\mathbf{z} \sim p_{noise}} [\log(1 - D(G(\mathbf{z}; \theta_G); \theta_D))] \quad (3.1)$$

其中 p_{real} 表示原始数据集服从的概率密度分布函数， $\mathbf{x}$ 为服从 p_{real} 的随机变量； p_{noise} 表示生成网络的输入噪声服从的概率密度分布函数， $\mathbf{z}$ 为服从 p_{noise} 的随机变量； $D(\mathbf{x}; \theta_D)$ 表示判别网络以 $\mathbf{x}$ 作为输入时的输出，其中 θ_D 为判别网络的网络参数； $G(\mathbf{z}; \theta_G)$ 表示生成网络以 $\mathbf{z}$ 作为输入时的输出，其中 θ_G 为生成网络的网络参数。判别网络的输出表示判别网络在当前网络参数下，判定当前输入来自于原始数据集的概率（即服从 p_{real} 的概率），生成网络的输出则表示生成网络在当前网络参数下，基于当前输入的生成数据。因此 $1 - D(G(\mathbf{z}; \theta_G); \theta_D)$ 实质上是判别网络正确判断出其输入数据来自于生成数据的概率，而 $V(D, G)$ 的实际意义为判别网络将其输入正确分类的概率。

对目标函数进行求解时，其内层优化是求解使 $V(D, G)$ 最大化的 θ_D ，即对判别网络的优化，其外层优化则是求解使 $V(D, G)$ 最小化的 θ_G ，即对生成网络的优化。对于内层优化而言，由于 θ_G 给定，因此目标函数可以转化为：

$$\begin{aligned}
 V(D, G) &= \mathbb{E}_{\mathbf{x} \sim p_{real}} [\log D(\mathbf{x}; \theta_D)] + \mathbb{E}_{\mathbf{z} \sim p_{noise}} [\log(1 - D(G(\mathbf{z}; \theta_G); \theta_D))] \\
 &= \mathbb{E}_{\mathbf{x} \sim p_{real}} [\log D(\mathbf{x}; \theta_D)] + \mathbb{E}_{\mathbf{x} \sim p_G} [\log(1 - D(\mathbf{x}; \theta_D))] \\
 &= \int p_{real}(\mathbf{x}) \log D(\mathbf{x}; \theta_D) + p_G(\mathbf{x}) \log(1 - D(\mathbf{x}; \theta_D)) d\mathbf{x}
 \end{aligned} \tag{3.2}$$

其中 p_G 为生成网络输出服从的概率密度分布函数。通过求导可得对于给定的生成网络 G ，判别网络的最优解 $D_G^*(\mathbf{x})$ 为：

$$D_G^*(\mathbf{x}) = \frac{p_{real}(\mathbf{x})}{p_{real}(\mathbf{x}) + p_G(\mathbf{x})} \tag{3.3}$$

基于 $D_G^*(\mathbf{x})$ ，外层优化问题可转换为：

$$\begin{aligned}
 \min_{\theta_G} V(G, D_G^*(\mathbf{x})) &= \int p_{real}(\mathbf{x}) \frac{p_{real}(\mathbf{x})}{p_{real}(\mathbf{x}) + p_G(\mathbf{x})} + p_G(\mathbf{x}) \frac{p_G(\mathbf{x})}{p_{real}(\mathbf{x}) + p_G(\mathbf{x})} d\mathbf{x} \\
 &= 2D_{JS}(p_{real} \parallel p_G) - \log 4
 \end{aligned} \tag{3.4}$$

其中 $D_{JS}(p_{real} \parallel p_G)$ 为 p_{real} 和 p_G 之间的 Jensen-Shannon (JS) 散度^[87]：

$$D_{JS}(p_{real} \parallel p_G) = \frac{1}{2} \int p_{real}(\mathbf{x}) \frac{2p_{real}(\mathbf{x})}{p_{real}(\mathbf{x}) + p_G(\mathbf{x})} + p_G(\mathbf{x}) \frac{2p_G(\mathbf{x})}{p_{real}(\mathbf{x}) + p_G(\mathbf{x})} d\mathbf{x} \tag{3.5}$$

由于两个概率密度分布函数之间的 JS 散度具有非负性，且当且仅当两个分布函数相等时其值为 0。因此可知当 $p_{real} = p_G$ 时外层优化问题取得最优解，即生成网络的输出服从 p_{real} 分布时。此时有 $D_G^*(\mathbf{x}) = \frac{1}{2}$ ，即判别网络再也无法区分其输入来自原始数据集或是生成网络的输出（因为与随机判别的概率相同）。因此在取得 GANs 的目标函数的最优解时，生成网络实现了对原始数据集服从的概率分布的学习。

3.3. 多风电场风速联合概率分布估计

如上节所述，优化问题(3.1)的最优解可以保证生成网络实现对原始数据集的概率密度分布的学习和完全复制，但是由于神经网络的复杂特性，直接求解(3.1)的难度较大，因此实际工程中往往采用梯度法进行求解。本节将以历史风电数据作为原始数据集，基于梯度法对 GANs 进行训练，在考虑多风电场之间的空间相

关性的前提下，实现对风速的联合概率分布估计。

3.3.1. 基于梯度法的GANs迭代训练算法

令 $\Omega = \{\mathbf{ws}_1, \mathbf{ws}_2, \dots, \mathbf{ws}_n\}$ 表示历史风电数据集，其中 n 表示数据集中数据的总数。具体地， Ω 中的元素均为 N_w 维向量：

$$\mathbf{ws}_i = (\mathbf{ws}_{i1}, \mathbf{ws}_{i2}, \dots, \mathbf{ws}_{iN_w}) \quad (3.6)$$

其中 N_w 为风电场的总数； $\mathbf{ws}_{ij}$ 表示在 i 时刻第 j 个风电场的风速。

生成网络的输入噪声 $\mathbf{Z}$ 为 M 维的随机向量，它的元素服从高斯分布，且满足独立同分布特性（Independent and identically distributed, IID）。

基于 GANs 的多风电场风速联合分布估计的具体算法如下：

步骤 1：从 Ω 中随机抽样 k 个数据，记为 $\{\mathbf{ws}_{(1)}, \mathbf{ws}_{(2)}, \dots, \mathbf{ws}_{(k)}\}$ ，按照 $\mathbf{Z}$ 的分布随机抽样 k 个数据，记为 $\{\mathbf{z}_{(1)}, \mathbf{z}_{(2)}, \dots, \mathbf{z}_{(k)}\}$ ；

步骤 2：基于批梯度上升法（Batch gradient ascent）^[75]对判别网络进行训练，对网络参数 θ_D 进行更新：

$$g_D = \nabla_{\theta_D} \frac{1}{k} \sum_{i=1}^k [\log D(\mathbf{ws}_{(i)}; \theta_D) + \log(1 - D(G(\mathbf{z}_{(i)}; \theta_G); \theta_D))] \quad (3.7)$$

$$\theta_D \leftarrow \theta_D + \eta_D g_D \quad (3.8)$$

其中 η_D 表示判别网络的学习率。对比式(3.1)和式(3.7)可以看出，式(3.7)实质上是基于随机采样的 $\{\mathbf{ws}_{(1)}, \mathbf{ws}_{(2)}, \dots, \mathbf{ws}_{(k)}\}$ 和 $\{\mathbf{z}_{(1)}, \mathbf{z}_{(2)}, \dots, \mathbf{z}_{(k)}\}$ 这两个各自包含 k 个数据的集合，通过求取算术平均值的方式对式(3.1)中包含期望值的目标函数进行估计，进而求出其对于 θ_D 的梯度 g_D ，并以此作为 θ_D 更新的方向。由于式(3.1)的内层优化是求取最大值，因此在此处采用梯度上升。这种基于批梯度的算法，比起传统的梯度算法，可以有效地避免求解陷入局部最优的问题^[75]。

步骤 3：再次按照 $\mathbf{Z}$ 的分布随机抽样 k 个数据，记为 $\{\mathbf{z}_{(1)}, \mathbf{z}_{(2)}, \dots, \mathbf{z}_{(k)}\}$ ；

步骤 4：基于批梯度下降法（Batch gradient descent）^[75]对生成网络进行训练，

对网络参数 θ_G 进行更新：

$$g_G = \nabla_{\theta_G} \frac{1}{k} \sum_{i=1}^k \log(1 - D(G(\mathbf{z}_{(i)}; \theta_G); \theta_D)) \quad (3.9)$$

$$\theta_G \leftarrow \theta_G - \eta_G g_G \quad (3.10)$$

其中 η_G 表示生成网络的学习率。与步骤 3 类似，式(3.9)实质上是基于 $\{\mathbf{z}_{(1)}, \mathbf{z}_{(2)}, \dots, \mathbf{z}_{(k)}\}$ ，通过求取算术平均值的方式对 $\mathbb{E}_{\mathbf{z} \sim p_{\text{noise}}} [\log(1 - D(G(\mathbf{z}; \theta_G); \theta_D))]$ 进行估计，进而求出其对于 θ_G 的梯度 g_G ，并以此作为 θ_G 更新的方向。由于式(3.1)的外层优化是求取最小值，因此在此处采用梯度下降。同时，由于生成网络的训练目标是最小化判别网络正确分类生成网络的输出的概率，因此式(3.9)与式(3.7)对比，仅包含生成网络的输出部分而不包含历史风电数据部分。

步骤 5：如果已经达到最大训练次数，则结束对 GANs 的训练。否则返回步骤 1 开始下一轮训练。

3.3.2. 相关性分析

通过训练，GANs 将实现对数据集 Ω 分布的估计，具体表现为生成网络可以生成逼近 Ω 的分布的数据集。由上述步骤 1-5 可以看出，生成网络的训练实质是学习并建立一个从 $\mathbf{Z}$ 服从的分布（高斯分布）到 Ω 服从的分布（风速联合概率分布）的映射。在 GANs 的训练过程中，不同风电场风速之间的空间相关性也将被自动估计并包含在 GANs 的概率分布估计结果中。这里所谓的空间相关性，是指在同一时刻，位于同一区域不同地理位置的风电场风速不相互独立，而是存在一定相关性。如图 3-2 所示，由于同一时刻不同风电场的风速被封装成向量 $\mathbf{ws}_{(i)}$ 的形式一并馈入 GANs，因此空间相关性存在于 $\mathbf{ws}_{(i)}$ 的每一个元素之间。由于生成网络的输出服从的分布是对风速联合概率分布的近似，因此生成网络的输出数据也将具有同样的空间相关性，从而实现对联合概率分布的估计。经过训练后的生成网络，可以视为一个风速数据生成器，其生成的数据所满足的分布，就是对多风电场风速联合概率分布的估计。

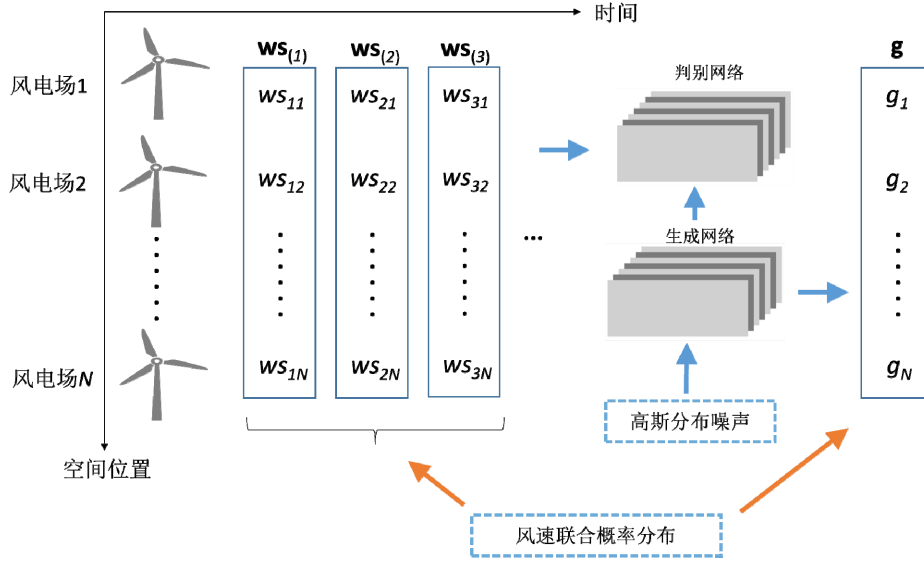


图 3-2 GANs 估计风速相关性示意图

3.4. 算例分析

本节将使用真实风速数据对 GANs 进行训练，重点关注单个风电场风速的边缘概率分布与多风电场风速联合概率分布的估计精度，以验证本文提出方法的有效性。

3.4.1. 数据与参数概况

本文从同一地区选取 8 个相互邻近的风电场进行历史风速数据采集。其中 70% 的数据作为 GANs 的训练数据，30% 的数据作为验证数据集。以现有文献的经验为基础^[88]，我们对 GANs 的超参数进行如下设定。生成网络是一个 8 层的多层感知机（Multi-layer perceptron, MLP），为了避免在神经网络训练中常见的梯度饱和以及消失问题，将其激活函数设定为带泄露的修正线性单元（Rectified linear unit, ReLU）^[75]：

$$f_{LR}(x) = \begin{cases} x, & x \geq 0 \\ 0.3x, & x < 0 \end{cases} \quad (3.11)$$

生成网络输出层的维度为 8，与风电场总数一致。判别网络则是一个 7 层的 MLP，其输出层激活函数是一个 sigmoid 单元^[75]：

$$f_s(x) = \frac{1}{1 + e^{-x}} \quad (3.12)$$

由于 **sigmoid** 单元的输出范围为 0 到 1 之间，因此可以用来表示概率值。与生成网络相同，判别网络的其它层激活函数均采用带泄露的 **ReLU**。生成网络的输入随机噪声是一个 4 维的 **IID** 高斯向量。采用 0.01 的学习率对 **GANs** 进行训练。

3.4.2. 风速概率分布估计结果

本节将对利用 **GANs** 对多风电场风速的联合概率分布估计结果进行展示和分析，同时我们也采用了目前常用的 **Copula** 算法进行概率分布估计作为对比，以验证我们提出方法的性能。

(1) Copula 算法

Copula 算法可以实现对多维随机变量的联合概率分布的估计，在电力系统新能源领域获得广泛的应用。**Copula** 算法的主要思路是先对每个随机变量的边缘概率分布进行估计，然后基于多个边缘概率分布计算其联合概率分布。根据实践经验和历史数据分析，基于 **Weibull** 分布进行单个风电场风速概率分布估计，是目前最常用的估计方法之一^[45]。服从 **Weibull** 分布的随机变量，其概率密度分布函数为：

$$f_w(x; \lambda, k, d) = \begin{cases} \frac{k}{\lambda} \left(\frac{x-d}{\lambda} \right)^{k-1} e^{-\left(\frac{x-d}{\lambda} \right)^k}, & x \geq d \\ 0, & x < d \end{cases} \quad (3.13)$$

其中 k 为形状参数（Shape parameter）， λ 为比例参数（Scale parameter）， d 为阈值参数（Threshold parameter）。对单个风电场风速的概率分布进行估计，就是确定 k ， λ ， d 三个参数的值，可以通过 **MLE** 实现：

$$\begin{aligned} (k_j, \lambda_j, d_j) &= \arg \max_{(k, \lambda, d)} \prod_{i=1}^n f_w(ws_{ij}; \lambda, k, d) \\ &= \arg \max_{(k, \lambda, d)} \sum_{i=1}^n \log f_w(ws_{ij}; \lambda, k, d) \end{aligned} \quad (3.14)$$

其中 k_j, λ_j, d_j 是第 j 个风电场风速概率分布估计对应的参数。

在确定了单个风电场风速的概率分布函数后，利用 **Copula** 算法对它们的相关性进行估计。根据 **Sklar** 定理^[89]，多风电场风速的联合累计分布函数（Cumulative distribution function, CDF）可以表示为：

$$\begin{aligned}
 F(ws_1, ws_2, \dots, ws_{N_w}) &= C(F_1(ws_1), F_2(ws_2), \dots, F_N(ws_{N_w})) \\
 &= \Phi_M(\Phi^{-1}(F_1(ws_1)), \Phi^{-1}(F_2(ws_2)), \dots, \Phi^{-1}(F_N(ws_{N_w})); \boldsymbol{\mu}, \boldsymbol{\Sigma})
 \end{aligned} \tag{3.15}$$

其中 ws_j 表示第 j 个风电场的风速, F_j 为第 j 个风电场的风速的 CDF; C 为 Copula 函数, Φ^{-1} 为标准高斯分布的 CDF 的逆函数, $\Phi_M(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ 为 N 维高斯分布 $\varphi(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ 对应的 CDF:

$$\varphi(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{e^{\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)}}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}|}} \tag{3.16}$$

其中 $\boldsymbol{\mu}$ 为均值向量, 维度为 N , $\boldsymbol{\Sigma}$ 为协方差矩阵, 维度为 $N \times N$, $|\boldsymbol{\Sigma}|$ 为 $\boldsymbol{\Sigma}$ 的行列式。与式(3.14)类似, 可以通过 MLE 计算 $\boldsymbol{\mu}$ 和 $\boldsymbol{\Sigma}$ 的值:

$$\begin{aligned}
 (\boldsymbol{\mu}, \boldsymbol{\Sigma}) &= \arg \max_{(\boldsymbol{\mu}, \boldsymbol{\Sigma})} \prod_{i=1}^n C(F_1(ws_{i1}), F_2(ws_{i2}), \dots, F_N(ws_{iN_w})) \\
 &= \arg \max_{(\boldsymbol{\mu}, \boldsymbol{\Sigma})} \sum_{i=1}^n \log C(F_1(ws_{i1}), F_2(ws_{i2}), \dots, F_N(ws_{iN_w}))
 \end{aligned} \tag{3.17}$$

从而实现对多风电场风速的联合概率分布的估计。

(2) 概率分布估计结果

图 3-3 展示了 GANs 和 Copula 算法对单个风电场的概率分布的估计结果。结果显示 GANs 和 Copula 算法都能够实现对给定风速数据集的概率分布函数的大致估计, 挖掘真实分布的主要特征。但是在风速小于 5m/s 的区间, Copula 算法对概率密度的估计有偏低的现象, 尤其是在概率密度峰值附近较为显著, 而 GANs 在这一区间的估计准确性明显更高。在风速大于 5m/s 的区间, GANs 和 Copula 算法都能较好地实现对真实概率分布的拟合。

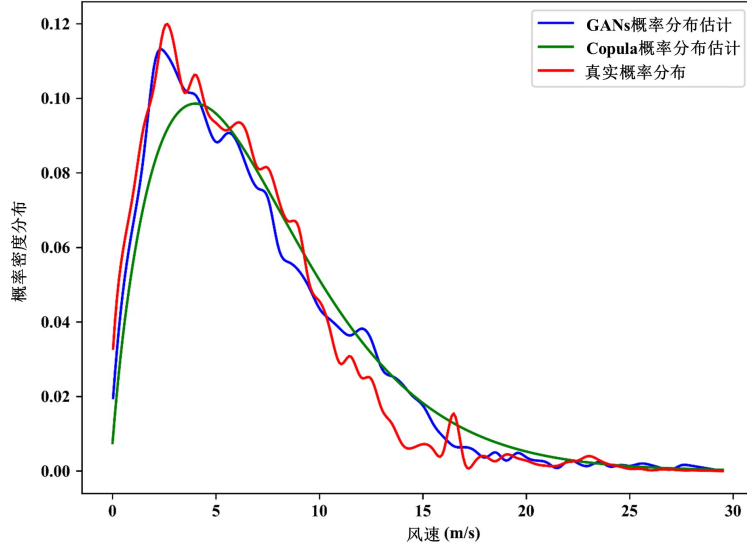


图 3-3 单个风电场风速概率分布结果估计

为了更加精确地展现和对比概率分布的估计结果,我们采用 Kullback-Leibler (KL)距离来量化地评估不同概率分布之间的相似程度。KL 距离通常被用来表示两个概率密度分布函数之间的距离,其表达式为:

$$D(p, q) = \int p(x) \log \frac{p(x)}{q(x)} dx \quad (3.18)$$

其中 $D(p, q)$ 即为分布 p 和 q 之间的 KL 距离。 $D(p, q)$ 具有非负性,且当 p 和 q 完全相等时,其值为 0,因此具有很好的距离特性。图 3-4 展示了对于每一个风电场,风电场风速估计概率与真实概率之间的 KL 距离。在 8 个风电场中,除了风电场 5,对于其它风电场,基于 GANs 的估计概率分布与真实分布之间的 KL 距离(以下称真实分布-GANs 距离),都小于基于 Copula 算法的估计概率分布与真实分布之间的 KL 距离(以下称真实分布-Copula 距离)。真实分布-GANs 距离在风电场 4 达到最大,为 1.64×10^{-2} ,而在风电场 6 达到最小,为 0.34×10^{-2} ;真实分布-Copula 距离在风电场 7 达到最大,为 2.79×10^{-2} ,而在风电场 5 达到最小,为 0.50×10^{-2} 。在风电场 2,真实分布-GANs 距离与真实分布-Copula 距离的差距达到最大,后者达到前者的 3.3 倍。通过上述数据分析,可以看出 GANs 对于单个风电场的概率分布估计准确性优于 Copula 算法。

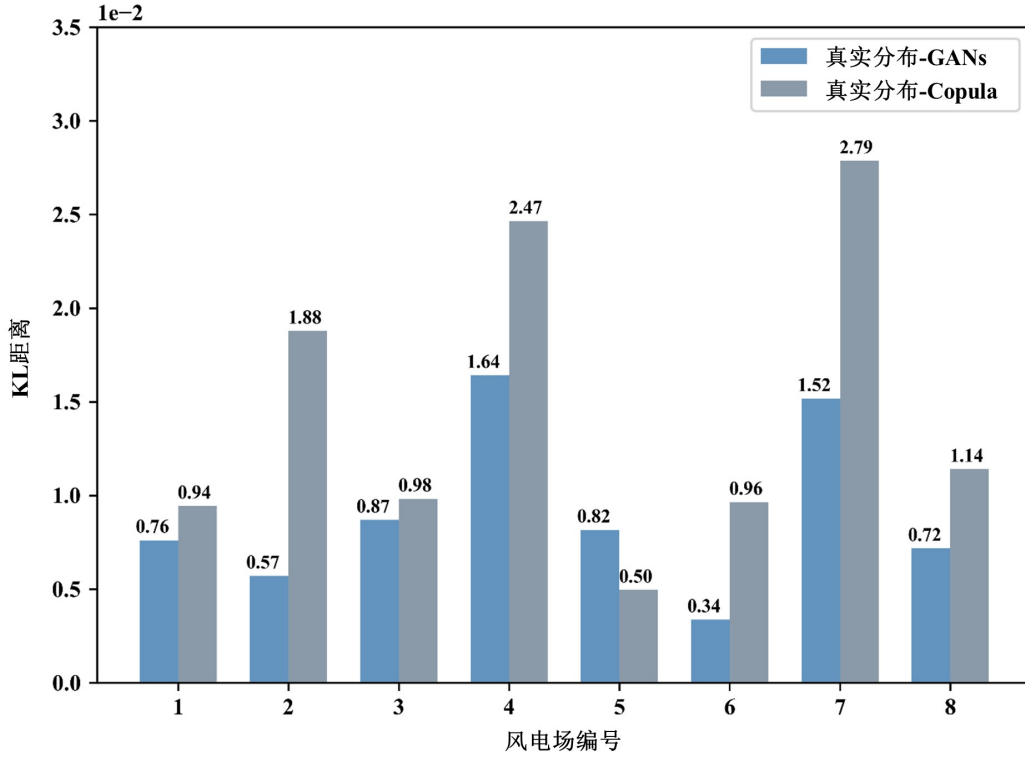


图 3-4 风电场风速估计概率与真实概率之间的 KL 距离

为了评估不同算法对多风电场风速之间的相关性估计准确性，我们引入了 Spearman 相关系数^[90]进行量化测算。Spearman 相关系数是在统计学中被广泛使用的相关性评估工具：

$$r(X_i, X_j) = \frac{Cov(P_i(X_i), P_j(X_j))}{\sqrt{Var(P_i(X_i)) \cdot Var(P_j(X_j))}} \quad (3.19)$$

其中 $r(X_i, X_j)$ 是随机变量 X_i 和 X_j 之间的 Spearman 相关系数， P_i 和 P_j 分别是 X_i 和 X_j 的 CDF； $Cov(X, Y)$ 表示 X 和 Y 之间的协方差， $Var(X)$ 表示 X 的方差。

Spearman 相关系数的取值范围为从 -1 到 1。Spearman 相关系数取值越接近 1，则表示两个变量的正相关性越强，越接近 -1 则表示两个变量的负相关性越强，越接近 0 则表示两个变量的相关性越弱。

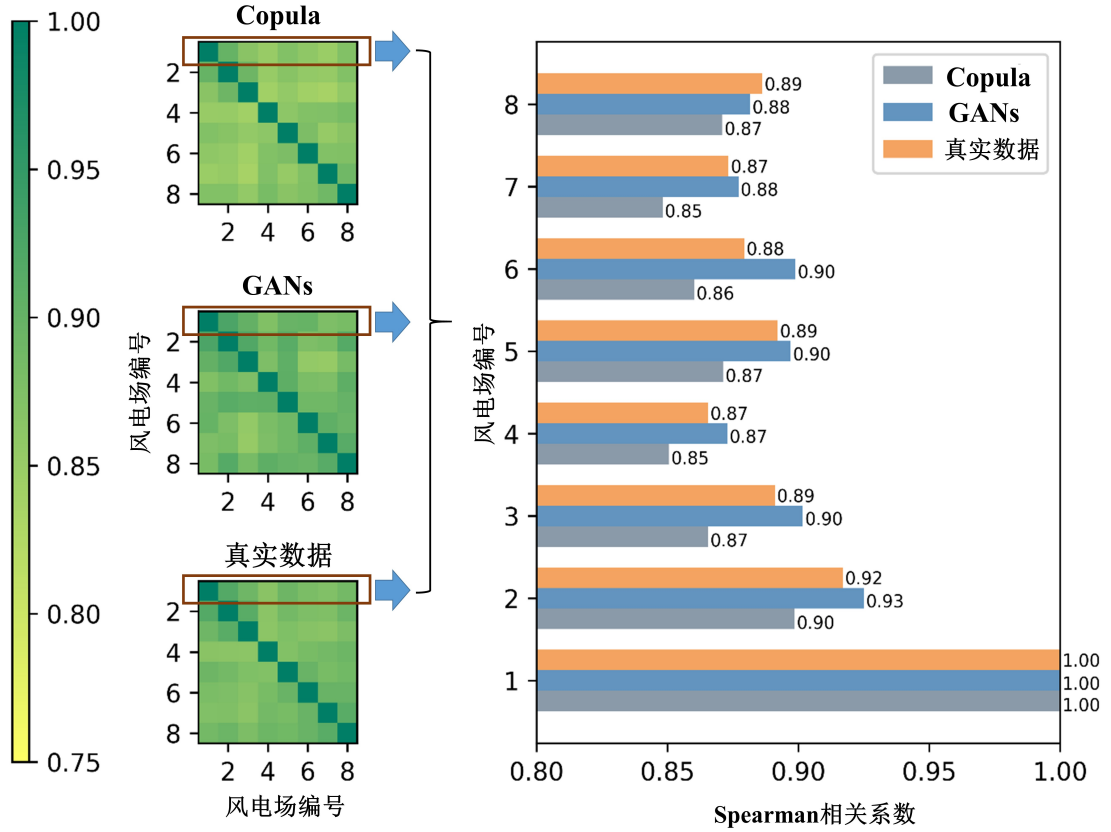


图 3-5 风电场风速相关性估计结果

在评估多个风电场风速之间的相关性时,由于需要考虑的随机变量总数大于 2, 因此可以采用 Spearman 相关系数矩阵进行计算。对于一组随机变量 $(X_1, X_2, \dots, X_{N_w})$, 它的 Spearman 相关系数矩阵 SM 维度是 $N_w \times N_w$, 具体表达式如下所示:

$$SM = \begin{pmatrix} sm_{11} & \dots & sm_{1N_w} \\ \vdots & \ddots & \vdots \\ sm_{N_w 1} & \dots & sm_{N_w N_w} \end{pmatrix} \quad (3.20)$$

$$sm_{ij} = r(X_i, X_j)$$

分别计算真实风速数据, 服从 GANs 估计的联合概率分布的数据, 以及服从 Copula 算法估计的联合概率分布的数据的 Spearman 相关系数矩阵, 其结果如图 3-5 所示。图 3-5 的左半部分, 通过热力图的形式展示了真实风速数据、GANs

估计和 Copula 算法估计对应的三个 Spearman 相关系数矩阵。从真实风速数据对应的 Spearman 相关系数矩阵可以看出, 矩阵中每一个元素取值都至少大于 0.75, 这表明本算例中的 8 个风电场的风速之间存在比较强的正相关性。此外, 从另外两个矩阵可以看出, GANs 和 Copula 算法的估计结果, 都包含了风电场风速之间的强正相关性, 这表明这两种算法都能够估计出不同随机变量间的相关性。为了进一步对比 GANs 和 Copula 算法对风电场风速相关性的估计准确性, 图 3-5 的右半部分选取了左半部分 3 个 Spearman 相关系数矩阵的第一行元素, 即风电场 1 和其它 7 个风电场风速之间的 Spearman 相关系数, 以柱状图的形式进行展示。可以看出 Copula 算法倾向于低估不同风电场风速之间的相关性 (所有 7 个 Spearman 相关系数都低于真实风速数据对应的 Spearman 相关系数), 而 GANs 倾向于高估不同风电场风速之间的相关性 (除风电场 8, 其余 6 个 Spearman 相关系数都高于真实风速数据对应的 Spearman 相关系数)。同时, 可以看出与 Copula 算法相比, 基于 GANs 估计的 Spearman 相关系数与真实风速数据更加接近。具体地, 除了在风电场 6 处 Copulas 算法的估计结果更加接近真实风速数据, 在其余所有风电场处 GANs 均取得更加接近真实数据的估计结果。

最后利用 L_2 距离计算 3 个 Spearman 相关系数矩阵的相似性, 进行进一步的量化比较:

$$L_2(\mathbf{X}, \mathbf{Y}) = \sqrt{\sum_i \sum_j (x_{ij} - y_{ij})^2} \quad (3.21)$$

其中 x_{ij} 和 y_{ij} 分别表示矩阵 $\mathbf{X}$ 和 $\mathbf{Y}$ 的第 i 行第 j 列元素。计算结果表明, 真实风速数据和 GANs 对应的 Spearman 相关系数矩阵之间的 L_2 距离为 0.109, 真实风速数据和 Copula 算法对应的 Spearman 相关系数矩阵之间的 L_2 距离为 0.156, 是前者的 1.43 倍。这进一步表明 GANs 对风电场风速相关性的估计准确性更高。

综上所述, GANs 和 Copula 算法均能实现对多风电场风速联合概率分布的估计, 但是 GANs 能够实现更加精确的估计结果。

3.5. 本章小结

为了测算风电出力不确定性对电力系统可靠性的影响, 本章对风速概率分布进行估计, 作为后续可靠性估计的基础。具体地, 本章利用历史风速数据对

GANs 进行训练，在不依赖任何先验知识的前提下，实现对多风电场风速概率分布的估计，同时考虑了同一区域处于不同位置的风电场风速之间的空间相关性。算例分析结果显示，GANs 在单个风电场风速概率分布估计以及多风电场风速相关性估计这两方面都取得比常用的 Copula 方法更加准确的结果。

4. 考虑供需侧不确定性的电力系统可靠性分析

4.1. 引言

电力系统可靠性分析是电力系统安全领域中的一个重要课题^[54]。伴随电力系统运行日益复杂化的趋势，电力系统可靠性分析评估的复杂度也逐渐提高。无论是电力系统中需求侧灵活资源参与需求侧响应，还是供给侧具有出力间歇性的新能源机组的并网，乃至系统中任一元件的运行状态，都会对电力系统的安全运行和可靠性造成直接影响^[14,34]。

在对复杂电力系统进行可靠性分析时，为了充分考虑前文所述的系统供需双侧的不确定性以及系统元件的运行状态，通常可采用蒙特卡洛模拟法（Monte Carlo simulation, MC）算法。但是对于复杂度较高的问题，MC 算法存在采样数量要求较高、收敛速度较慢的缺点，尤其是针对低概率事件的估计^[63]。为了提高计算效率，本章采用交叉熵重要度采样（Cross-entropy importance sampling, CE）算法，实现对电力系统可靠性的高效分析评估。然而，由于 CE 算法的使用要求被评估对象中的不确定变量的概率分布函数可以显式化及参数化地表示，因此在处理只有原始数据而概率分布未知的问题上存在较大障碍^[66]。本文进行分析评估的问题——考虑温控负荷需求侧响应和风电接入的电力系统可靠性分析，正是一个典型的无法直接使用 CE 算法的问题。为解决上述困难，本章基于前两章对供需双侧不确定性的处理，通过数据驱动方法将原电力系统可靠性分析问题，转换为可以直接采用 CE 算法进行评估的问题，再通过 CE 算法实现高效分析评估，最后通过算例分析验证所提方法的有效性。

4.2. 交叉熵重要度采样算法

CE 算法是一种具有较快收敛速度的采样算法，是一种针对 MC 算法的改进算法，尤其是针对低概率事件，对比 MC 算法在性能上有较显著提高^[91]。本节将对 CE 算法的基本原理进行介绍。

假定要对以下事件对应的概率进行估计：

$$h = P[S(X) \geq \gamma] \quad (4.1)$$

其中 X 为随机变量, $S(X)$ 为某一以 X 为自变量的值函数 (Value function), γ 为常数, 则 h 表示事件 “ $S(X)$ 取值大于或等于 γ ” 的概率。由于 h 可以通过数学期望的形式重新表示:

$$\begin{aligned} h &= \mathbb{E}_{X \sim f(;u)} I[S(X) \geq \gamma] \\ &= \int I[(S(X) \geq \gamma)] f(X; u) dX \end{aligned} \quad (4.2)$$

其中 $f(;u)$ 为 X 的概率密度分布函数, u 为唯一确定 f 所需的参数, I 为指示函数 (Indicator function):

$$I(\text{事件}k) = \begin{cases} 1, & \text{事件}k \text{为真} \\ 0, & \text{事件}k \text{为假} \end{cases} \quad (4.3)$$

因此可以采用 MC 算法对 h 进行如下估计:

$$\hat{h} = \frac{1}{N_m} \sum_{i=1}^{N_m} I[S(X_{(i)}) \geq \gamma] \quad (4.4)$$

其中 $\hat{h}$ 表示对 h 的近似, $X_{(i)}$ 表示基于 $f(;u)$ 进行采样所得的第 i 个数据样本, N_m 表示抽样总数量。根据大数定律, 随着 N_m 的增大, 基于 MC 算法所得的 $\hat{h}$ 取值将会趋于收敛至 h 。因此当 N_m 的数量足够大时, $\hat{h}$ 是对 h 的一个准确度较高的估计。

然而, 当 h 的值比较小时, 为了保证对 h 估计的准确, MC 算法对 N_m 的数量规模要求较高: 由于 $S(X) \geq \gamma$ 的概率远低于 $S(X) < \gamma$, 抽样样本中绝大部分数据都是满足 $S(X) < \gamma$, 为了使抽样样本中满足 $S(X) \geq \gamma$ 的数据规模足以满足大数定律, N_m 的取值需要很大, 这大大提高了评估所花费的时间。为了提高评估性能, 可采用重要度采样技巧^[91]:

$$\begin{aligned} h &= \int I[(S(X) \geq \gamma)] g(X) \frac{f(X; u)}{g(X)} dX \\ &= \mathbb{E}_{X \sim g} \left\{ I[(S(X) \geq \gamma)] \frac{f(X; u)}{g(X)} \right\} \end{aligned} \quad (4.5)$$

其中 g 为任意概率分布函数。显然，通过重要度采样技巧， h 被等价转换为一个基于新分布 g 的期望，因此 h 可以通过下式进行估计：

$$\hat{h} = \frac{1}{N_m} \sum_{i=1}^{N_m} I[S(X_{(i)}) \geq \gamma] W(X_{(i)}) \quad (4.6)$$

$$W(X) = \frac{f(X; u)}{g(X)} \quad (4.7)$$

其中 $W(X)$ 被称作似然比 (Likelihood ratio, LR)，即原概率分布函数与新概率分布函数之间的比值，式(4.6)中的 $X_{(i)}$ 为基于 g 进行采样所得的第 i 个数据样本。

如果我们将 g 构建为如下式所示：

$$g^*(X) = \frac{I[S(X) \geq \gamma] f(X; u)}{h} \quad (4.8)$$

则式(4.6)将变为：

$$\hat{h} = \frac{1}{N_m} \sum_{i=1}^{N_m} h \quad (4.9)$$

显然此时只需要 1 个数据样本，即 $N_s = 1$ ， $\hat{h}$ 便是 h 的一个无偏估计，因此将 $g^*(X)$ 成为重要度采样技巧中的最优概率分布函数，它可以实现采样数量的最小化。

然而上述算法存在的问题是， $g^*(X)$ 中包含了待估计量 h ，因此实际上在完成估计之前无法直接获取 $g^*(X)$ 。CE 算法的核心思想，就是求取 $g^*(X)$ 的近似值进而实现对 h 的高效估计。如章节 3.4.2 中所述，KL 距离可以用来评估两个概率分布函数之间的相似性，CE 算法同样基于 KL 距离评估其对 $g^*(X)$ 的近似效果。CE 算法的目标是求取参数 v 使得 KL 距离 $D(g^*, f(v))$ 最小化：

$$v^* = \arg \min_v \left(\int g^*(X) \log g^*(X) dX - \int g^*(X) \log f(X; v) dX \right) \quad (4.10)$$

显然，这等效于：

$$\begin{aligned} v^* &= \arg \max_v \int g^*(X) \log f(X; v) dX \\ &= \arg \max_v \mathbb{E}_{X \sim f(u)} I[S(X) \geq \gamma] \log f(X; v) \end{aligned} \quad (4.11)$$

引入新的概率分布函数 $f(\cdot; w)$ ，再次对式(4.11)采用重要度采样技术可得：

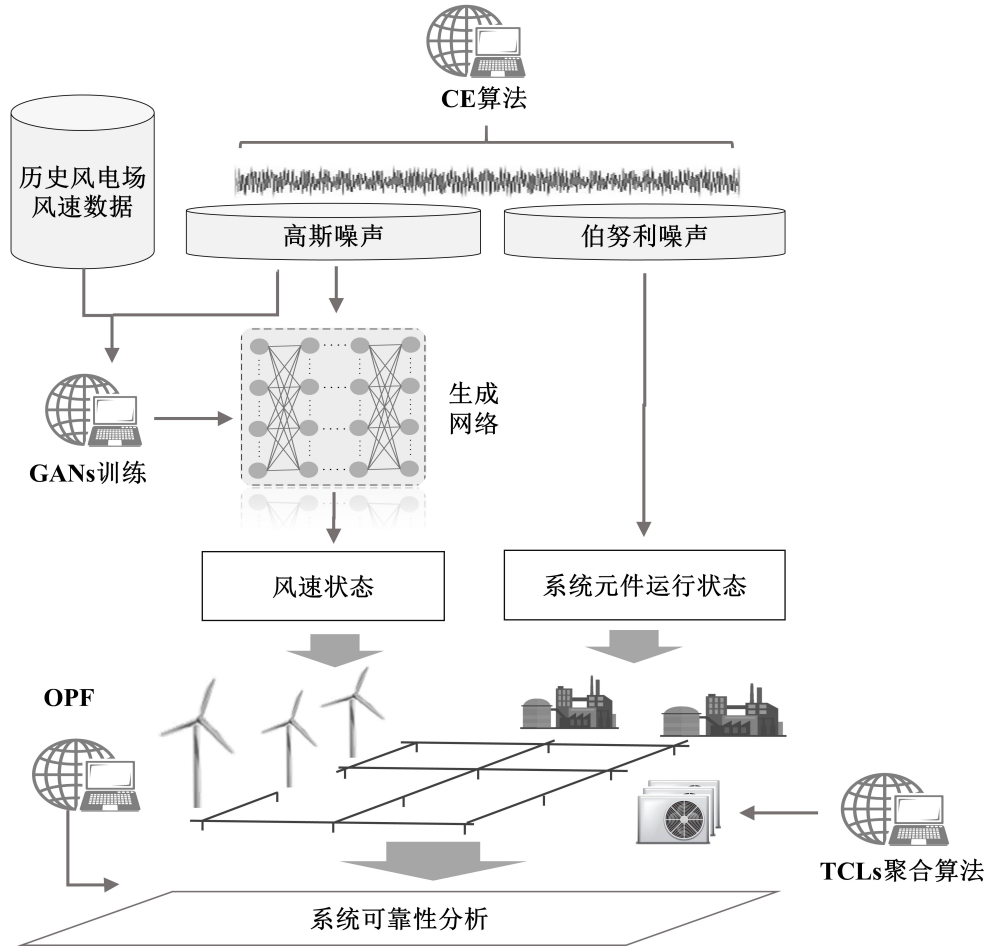
$$v^* = \arg \max_v \mathbb{E}_{X \sim f(\cdot; w)} I[S(X) \geq \gamma] W(X; u, w) \log f(X; v) \quad (4.12)$$

其中将 w 称为参考参数， $W(X; u, w)$ 的表达式为：

$$W(X; u, w) = \frac{f(X; u)}{f(X; w)} \quad (4.13)$$

可见 CE 算法本质上是一个关于 v 的优化问题，下一节将结合电力系统可靠性问题对其求解算法进行讲解。

4.3. 考虑供需双侧不确定性的电力系统可靠性分析



基于前文对供需双侧不确定性的数据驱动估计算法，本节将基于如图 4-1 所示的流程，开展电力系统可靠性的高效分析评估。通过历史风电场风速数据训练

GANs, 基于 CE 算法, 利用高斯噪声和伯努利噪声分别实现对风速和系统元件运行状态的仿真, 基于 TCLs 聚合算法实现对 TCLs 集群参与需求侧响应备用容量的估计, 从而最终确定电力系统运行状态, 并通过求解最优潮流 (Optimal power flow, OPF) 完成电力系统可靠性分析。

4.3.1. 问题构建

本文进行可靠性分析的电力系统, 包含位于不同位置的多个风电场, 并考虑 TCLs 参与需求侧响应提供的运行备用。实质上基于第 2 章的 TCLs 聚合算法, 可以将聚合 TCLs 等效为具有备用的发电机组。为了确定系统的运行状态, 我们引入 $\mathbf{ws}$ 和 $\mathbf{A}$ 这两个状态向量, 其中 $\mathbf{ws}$ 为 N_w 维风电场风速向量, 其元素为各个风电场对应风速; $\mathbf{A}$ 为系统元件运行状态向量:

$$\begin{aligned} \mathbf{A} &= (A_1, A_2, \dots, A_C) \\ A_i &= \begin{cases} 0, & \text{故障状态} \\ 1, & \text{正常状态} \end{cases} \end{aligned} \quad (4.14)$$

其中 A_i 表示第 i 个系统元件的运行状态, C 表示系统中元件的总数, 包含发电机组和输电线路。

我们通过以下两个指标对电力系统的可靠性进行分析评估: 负荷损失概率 (Loss of load probability, LOLP) 和功率供应不足期望 (Expected power not served, EPNS) [92]。其中 LOLP 表征电力系统供电不足的概率, EPNS 表征电力系统供电缺额功率的期望, 其表达式分别为:

$$LOLP = \mathbb{E}_{\mathbf{ws}, \mathbf{A}} I[LC(\mathbf{ws}, \mathbf{A}) > \gamma] \quad (4.15)$$

$$EPNS = \mathbb{E}_{\mathbf{ws}, \mathbf{A}} LC(\mathbf{ws}, \mathbf{A}) \quad (4.16)$$

其中 LC 表示在 $\mathbf{ws}$ 和 $\mathbf{A}$ 确定的运行状态下系统的负荷切除功率。通常 γ 的取值为接近 0 的常数, 用于过滤系统小扰动或计算误差的干扰。显然, 为了求取可靠性指标 LOLP 和 EPNS, 首先必须求取不同 $\mathbf{ws}$ 和 $\mathbf{A}$ 对应的 LC 。

在给定 $\mathbf{ws}$ 和 $\mathbf{A}$ 的情况下, 电力系统的负荷切除功率 LC 可以通过求解以下直流 OPF 模型来获取:

$$\min_{\boldsymbol{\theta}, \mathbf{P}_g, \mathbf{P}_w, \mathbf{P}_T} LC = \mathbf{e}^T \mathbf{P}_{LC} \quad (4.17)$$

$$\text{s.t.} \quad \mathbf{B}_{bus} \boldsymbol{\theta} + \mathbf{P}_d - \mathbf{P}_{LC} - \mathbf{C}_g \mathbf{P}_g - \mathbf{C}_w \mathbf{P}_w - \mathbf{C}_T \mathbf{P}_T = 0 \quad (4.18)$$

$$|\mathbf{B}_L (\boldsymbol{\theta}_f - \boldsymbol{\theta}_t)| \leq \mathbf{F}_{\max} \quad (4.19)$$

$$\mathbf{P}_g^{\min} \leq \mathbf{P}_g \leq \mathbf{P}_g^{\max} \quad (4.20)$$

$$0 \leq \mathbf{P}_w \leq \mathbf{P}_w^{\max} \quad (4.21)$$

$$0 \leq \mathbf{P}_T \leq \mathbf{P}_T^{\max} \quad (4.22)$$

$$0 \leq \mathbf{P}_{LC} \leq \mathbf{P}_{LC}^{\max} \quad (4.23)$$

其中 $\boldsymbol{\theta}$ 表示系统节点电压相角向量，其维度等于系统节点数 N_{bus} ； $\mathbf{P}_g$ 为火电发电机组出力功率向量，其维度等于系统火电发电机组总数 N_g ； $\mathbf{P}_w$ 为风电场出力功率向量，其维度等于 N_w ； $\mathbf{P}_T$ 为 TCLs 参与需求侧响应功率向量，其维度等于系统 TCLs 集群总数 N_T ； $\mathbf{P}_d$ 和 $\mathbf{P}_{LC}$ 分别表示系统节点负荷功率向量和系统节点负荷切除功率向量，它们的维度均为 N_{bus} ； $\mathbf{e}$ 表示维度为 N_{bus} 的单位向量，它的所有元素均为 1； $\mathbf{B}_{bus}$ 为系统节点电纳矩阵，其维度为 $N_{bus} \times N_{bus}$ ； $\mathbf{B}_L$ 为系统线路电纳矩阵，其维度为 $N_L \times N_L$ （ N_L 为系统线路总数）； $\mathbf{C}_g$ 为 $N_{bus} \times N_g$ 的火电机组联接矩阵，它的所有元素取值为 0 或者 1，其中第 i 行第 j 列元素取 1 则表示第 j 台火电发电机组与第 i 个节点相连，取 0 则表示不相连； $\mathbf{C}_w$ 和 $\mathbf{C}_T$ 分别为 $N_{bus} \times N_w$ 的风电场联接矩阵和 $N_T \times N_w$ 的 TCLs 集群联接矩阵，其取值与 $\mathbf{C}_g$ 类似； $\boldsymbol{\theta}_f$ 和 $\boldsymbol{\theta}_t$ 分别为线路送端（From-end）节点电压相角向量和受端（To-end）节点电压相角向量，它们的维度均为 N_L ； $\mathbf{F}_{\max}$ 是 N_L 维的线路传输的热稳定极限功率向量； $\mathbf{P}_g^{\min}$ 和 $\mathbf{P}_g^{\max}$ 分别是 N_g 维的火电发电机组出力功率下限向量和上限向量； $\mathbf{P}_w^{\max}$ ， $\mathbf{P}_T^{\max}$ 和 $\mathbf{P}_{LC}^{\max}$ 分别是 N_w 维的风电出力功率上限向量， N_T 维的 TCLs 集群参

与需求侧响应功率上限向量，和 N_{bus} 维的节点负荷切除功率上限向量。

由于本文的可靠性分析问题中考虑系统元件的运行状态，因此在求解直流 OPF 模型时需要对应地对故障元件进行修正。对于处于故障状态的发电机组，上述直流 OPF 模型中的对应发电功率上限和下限将被设定为 0。同样的，对于处于故障状态的线路，直流 OPF 模型中对应的线路传输热稳定极限功率也将被设定为 0。

为了求解上述直流 OPF 模型，还需将风速转换为对应的风电出力功率。风电出力功率上限与风速之间的关系，可以通过以下分段模型给出^[93]：

$$P_w(V) = \begin{cases} 0, & V \leq V_{ci} \\ P_r \frac{V^2 - V_{ci}^2}{V_r^2 - V_{ci}^2}, & V_{ci} < V < V_r \\ P_r, & V_r \leq V \leq V_{co} \\ 0, & V > V_{co} \end{cases} \quad (4.24)$$

其中 $P_w(V)$ 表示风速 V 对应的风电出力功率； P_r 和 V_r 分别表示额定风电功率和额定风速； V_{ci} 和 V_{co} 分别表示切入风速和切出风速。

4.3.2. 可靠性分析

如上节所述，通过求解直流 OPF 模型(4.17)—(4.23)，可以获取负荷切除功率 LC 。而为了进一步实现对 LOLP 和 EPNS 的计算，由于 LOLP 和 EPNS 实质上都是某一随机事件的数学期望，因此如 4.2 节所述，可以通过基于特定的概率分布函数采样获取一定规模的数据样本并求取平均值的方式进行估计。本节将通过 CE 算法实现对 LOLP 和 EPNS 的高效计算。

如式(4.12)所示，CE 算法实质上是在显式参数化分布族 $\{f(\cdot; \mathbf{v})\}$ 中寻找对最优概率分布函数 $g^*(X)$ 的近似分布函数。所谓显式参数化是指函数 f 的形式可以通过公式明确表达，并且可以通过参数 $\mathbf{v}$ 唯一确定其形式。因此，正如引言部分所述，对于包含 TCLs 和风电场的电力系统，CE 算法难以被直接用于进行可靠性分析，原因在于 TCLs 的参数分布以及风电出力的分布均是未知的和非参数化的，而只具有相关数据集。对于 TCLs，我们通过聚合算法将其等效为具有备用的发

电机组，从而解决参数分布未知的问题。对于风电出力，由于 GANs 的生成网络是一个从输入高斯分布噪声到实际风速分布的一个映射，因此可以利用 GANs 的这一特性解决使用 CE 算法的困难。具体地，当对风电场风速 $\mathbf{ws}$ 进行采样时，首先采用 CE 算法对 M 维的高斯噪声 $\mathbf{n}$ 进行采样，然后再通过生成网络将其映射为风电场风速。因此 CE 算法实际上只需要关注显示参数化分布的高斯噪声，而无需关注风电场风速的隐式分布函数。如图 4-2 所示，利用 GANs 将对 $(\mathbf{ws}, \mathbf{A})$ 的采样转换为对 $(\mathbf{n}, \mathbf{A})$ 的采样，从而消除了 CE 算法应用于电力系统可靠性分析的困难。以下将给出可靠性分析的具体算法和流程。

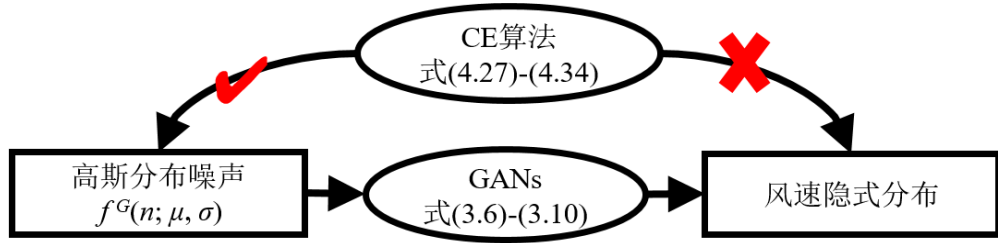


图 4-2 基于 CE 算法和 GANs 处理风速隐式分布问题示意图

由式(4.14)可知，系统元件状态 $\mathbf{A}$ 是一个伯努利变量，假设其对应分布参数为 $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_C\}$ ，则 $\mathbf{A}$ 中元素 A_i 的概率分布函数 $f^B(A_i; \lambda_i)$ 为：

$$f^B(A_i; \lambda_i) = (1 - \lambda_i)^{A_i} \lambda_i^{1-A_i} \quad (4.25)$$

并称 $f^B(\cdot; \lambda)$ 为 $\mathbf{A}$ 的概率分布函数。同样假设 $\mathbf{n}$ 对应的分布参数为 $\mu = \{\mu_1, \mu_2, \dots, \mu_M\}$ 和 $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_M\}$ ，则 $\mathbf{n}$ 中元素 n_i 的概率分布函数 $f^G(n_i; \mu_i, \sigma_i)$ 为：

$$f^G(n_i; \mu_i, \sigma_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(n_i - \mu_i)^2}{2\sigma_i^2}} \quad (4.26)$$

并称 $f^G(\cdot; \mu, \sigma)$ 为 $\mathbf{n}$ 的概率分布函数。

CE 算法的任务是根据(4.12)基于 KL 距离寻找参考参数 $\underline{\lambda}$ ， $\underline{\mu}$ 和 $\underline{\sigma}$ ，从而确定重要度采样的最优概率分布函数。由于直接求解该优化问题有较高的难度，因

此本文采用迭代算法进行求解。迭代算法的核心思想是通过迭代形成参考参数序列，不断扩充序列最终逼近最优解，具体如下^[91]：

步骤 1：算法初始化，令 $\lambda^0 = \lambda$ ， $\mu^0 = \mu$ ， $\sigma^0 = \sigma$ ，其中 λ^0 ， μ^0 ， σ^0 表示参考参数序列的初始值，令 $u = \{\lambda, \mu, \sigma\}$ 表示原始分布参数集。

步骤 2：每一轮迭代都将更新参考参数序列中的值，假设 $v^{t-1} = \{\lambda^{t-1}, \mu^{t-1}, \sigma^{t-1}\}$ 表示第 $t-1$ 个迭代伦次中更新的参考参数序列值。基于概率分布函数 $f^G(; \mu^{t-1}, \sigma^{t-1})$ 采样 k 个样本 $\{n_{(1)}, n_{(2)}, \dots, n_{(k)}\}$ ，基于概率分布函数 $f^B(; \lambda^{t-1})$ 采样 k 个样本 $\{A_{(1)}, A_{(2)}, \dots, A_{(k)}\}$ 。利用样本中的每一对 (n, A) 求解直流 OPF 模型获取负荷切除功率，并形成性能函数（Performance function）序列：

$$S = \{S_{(1)}, S_{(2)}, \dots, S_{(k)}\} \quad (4.27)$$

$$S_{(i)} = LC \left[G(n_{(i)}; \theta_G), A_{(i)} \right]$$

其中 $G(n_{(i)}; \theta_G)$ 表示训练完成后 GANs 的生成网络基于网络参数 θ_G 并以 $n_{(i)}$ 为输入噪声的输出。对 S 序列进行升序排序并以 S' 表示，并令 γ_t 表示序列 S' 的 $(1-\rho)$ 分位数，即序列 S' 中的第 $(1-\rho)k$ 个元素：

$$S' = \{S'_{(1)}, S'_{(2)}, \dots, S'_{(k)}\} \quad (4.28)$$

$$\gamma_t = S'_{((1-\rho)k)}$$

如果 γ_t 大于 γ ，则令 $\gamma_t = \gamma$ 。

步骤 3：通过求解以下问题计算 $v^t = \{\lambda^t, \mu^t, \sigma^t\}$ ：

$$\max_{v^t} \frac{1}{k} \sum_{i=1}^k I(S_{(i)} > \gamma_t) W(n_{(i)}, A_{(i)}; u, v^{t-1}) \log f(n_{(i)}, A_{(i)}; v^t) \quad (4.29)$$

其中 $W()$ 和 $f()$ 可以通过下式进行计算：

$$f(n, A; v^t) = \prod_{i=1}^C f^B(A_i; \lambda_i^t) \prod_{i=1}^M f^G(n_i; \mu_i^t, \sigma_i^t) \quad (4.30)$$

$$W(n, A; u, v^t) = \frac{f(n, A; u)}{f(n, A; v^t)} \quad (4.31)$$

对于伯努利分布和高斯分布，我们可以通过计算直接给出式(4.29)表示的优化问题的解析解：

$$\lambda_i^t = \frac{\sum_{j=1}^k I(S_{(j)} > \gamma_t) W(\mathbf{n}_{(j)}, \mathbf{A}_{(j)}; \mathbf{u}, \mathbf{v}^{t-1}) I(A_{(j)i} = 0)}{\sum_{j=1}^k I(S_{(j)} > \gamma_t) W(\mathbf{n}_{(j)}, \mathbf{A}_{(j)}; \mathbf{u}, \mathbf{v}^{t-1})} \quad (4.32)$$

$$\mu_i^t = \frac{\sum_{j=1}^k I(S_{(j)} > \gamma_t) W(\mathbf{n}_{(j)}, \mathbf{A}_{(j)}; \mathbf{u}, \mathbf{v}^{t-1}) n_{(j)i}}{\sum_{j=1}^k I(S_{(j)} > \gamma_t) W(\mathbf{n}_{(j)}, \mathbf{A}_{(j)}; \mathbf{u}, \mathbf{v}^{t-1})} \quad (4.33)$$

$$\sigma_i^t = \sqrt{\frac{\sum_{j=1}^k I(S_{(j)} > \gamma_t) W(\mathbf{n}_{(j)}, \mathbf{A}_{(j)}; \mathbf{u}, \mathbf{v}^{t-1}) (n_{(j)i} - \mu_i^{t-1})^2}{\sum_{j=1}^k I(S_{(j)} > \gamma_t) W(\mathbf{n}_{(j)}, \mathbf{A}_{(j)}; \mathbf{u}, \mathbf{v}^{t-1})}} \quad (4.34)$$

其中 $A_{(j)i}$ 和 $n_{(j)i}$ 分别表示 $\mathbf{A}_{(j)}$ 的第 i 个元素和 $\mathbf{n}_{(j)}$ 的第 i 个元素。

步骤 4: 如果 γ_t 小于 γ ，则令 $t \leftarrow t + 1$ 并重复步骤 2 进入下一轮迭代，否则认为最优参考参数为 $\mathbf{v} = \{\lambda^t, \mu^t, \sigma^t\}$ ，结束迭代算法。

通常， ρ 的取值为一个不是非常小的数，例如 0.01。根据式(4.28)，基于由最优参考参数确定的概率分布函数，可以使得性能函数不小于 γ 的概率不再非常低（ ρ 左右），从而降低抽样算法收敛所需的样本数量，提高算法的计算性能。根据最优参考参数 $\mathbf{v}$ ，LOLP 和 EPNS 可以通过下式进行估计：

$$LOLP = \frac{1}{N_r} \sum_{i=1}^{N_r} I \left\{ LC \left[G(\mathbf{n}_{(i)}; \theta_G), \mathbf{A}_{(i)} \right] \right\} W(\mathbf{n}_{(i)}, \mathbf{A}_{(i)}; \mathbf{u}, \mathbf{v}) \quad (4.35)$$

$$EPNS = \frac{1}{N_r} \sum_{i=1}^{N_r} LC \left[G(\mathbf{n}_{(i)}; \theta_G), \mathbf{A}_{(i)} \right] W(\mathbf{n}_{(i)}, \mathbf{A}_{(i)}; \mathbf{u}, \mathbf{v}) \quad (4.36)$$

其中 $(\mathbf{n}_{(i)}, \mathbf{A}_{(i)})$ 表示基于由 $\mathbf{v}$ 确定的概率分布函数抽样得到的第 i 组样本数据， N_r 表示抽样样本总数。

4.4. 算例分析

4.4.1. 系统概述

本节采用修改过的 RTS 测试系统^[94]作为算例进行可靠性分析评估以验证所提方法的有效性。系统的具体拓扑结构如图 4-3 所示。系统中一共有 24 台火电发电机组，分别分布在节点 1-15 之间，总装机容量为 2285MW。8 个风电场分别分布在节点 16-23 之间，总装机容量为 1840MW，它们的风速数据和分布估计采用第 3 章的算例分析结果。节点 15 处分布有第二章中的 TCLs 集群，通过参与需求侧响应为系统系统备用。系统元件的运行故障停用率参照文献[69]。

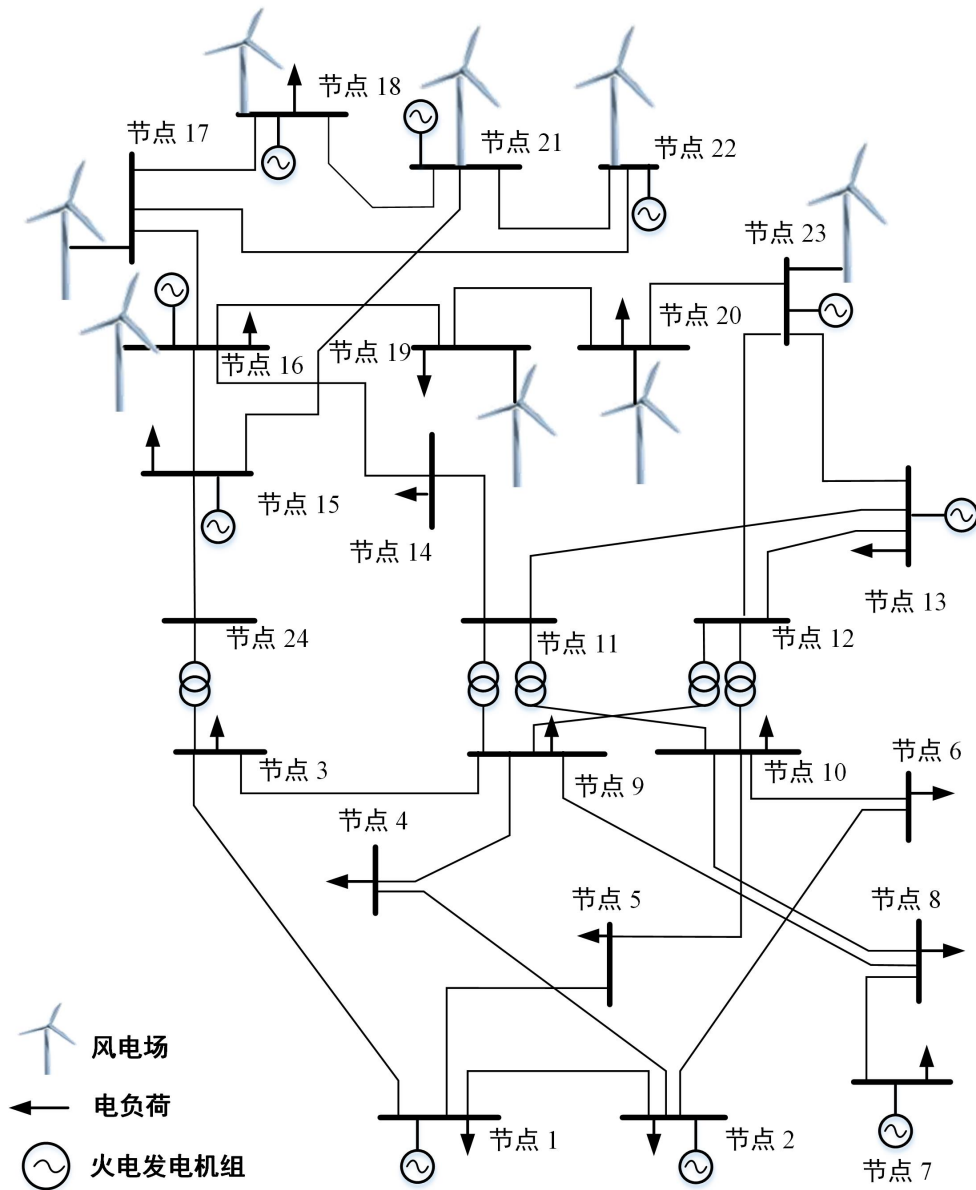


图 4-3 算例系统拓扑图

4.4.2. 可靠性分析结果

对上述电力系统进行可靠性分析，结果如表 4-1 所示。通过表 4-1 对 CE 评估算法和 MC 评估算法进行对比，从评估数值结果的角度可以看出，两种方法对 LOLP 和 EENS 这两个可靠性指标的计算结果均具有比较高的相似度。然而从计算效率的角度上看，两种算法收敛所需的采样数量具有较大差距。本文参考文献 [95] 采用以下收敛判据：

$$\frac{\sqrt{\text{Var}(\text{LOLP}(s))}}{\text{LOLP}(s)} \leq 0.01 \quad (4.37)$$

其中 $\text{LOLP}(s)$ 表示在采样数量 s 下的 LOLP 估计结果。由表 4-1 可知，CE 算法的收敛性能明显优于 MC 算法：CE 算法仅需要 58400 的采样数量即可达到收敛，而 MC 算法需要 246100 的采样数量才能实现收敛，这是 CE 算法所需采样数量的约 4.2 倍。这表明在本文所提的评估方法中，CE 算法的采用显著提高了计算效率。

表 4-1 系统可靠性指标计算结果

计算方法	LOLP/ 10^{-3}	EPNS/ 10^{-3}MW	采样数量
CE 评估算法	9.412	141.784	58400
MC 评估算法	9.256	139.504	246100

为了评估风电场装机容量对电力系统可靠性的影响，我们改变风电场容量在总发电容量中的占比，并分别计算对应情况下的可靠性指标。具体地，我们保持系统总发电容量不变，然后在增大风电场装机容量的同时等量地降低火电发电机组地发电容量。图 4-4 给出了在风电场装机容量增大一定数值下电力系统可靠性指标的变化，左右两幅图片分别表示 LOLP 和 EPNS 的变化。可以看出，随着风电场装机容量的增大，LOLP 和 EENS 都随之增大。这是因为在电力系统总发电容量为常数的前提下，风电场装机容量的增大导致了电力系统中具有天然出力波动性的发电份额的提高，从而增大了电力系统发电功率的不确定性，降低了供电质量，导致了电力系统的可靠性下降。

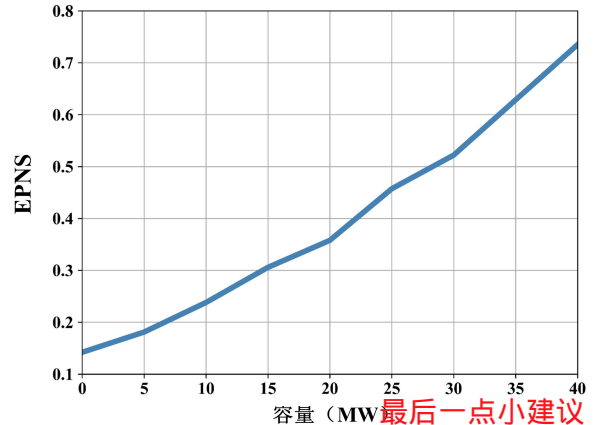
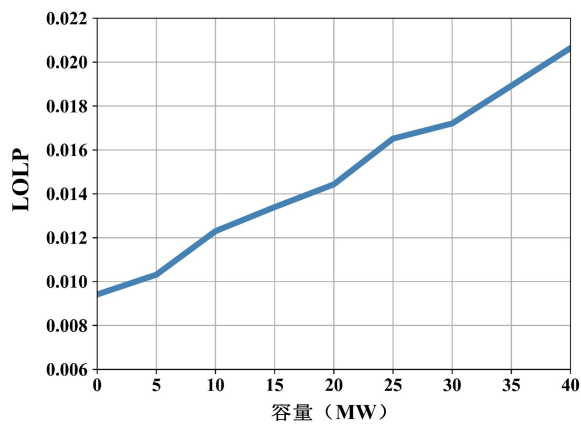


图 4-4 风电场容量对系统可靠性影响示意图

最后一点小建议：作为最后一章的算例，稍微有点单薄。只有两个图，4-4和4-5，分别说风电和温控负荷越多越好，有没有其他更吸引眼球的结论呢？

同时，为了评估 TCLs 集群参与需求侧响应对接入风电的电力系统可靠性的影响，我们改变 TCLs 备用容量并计算对应情况下的可靠性指标。具体地，我们一方面调整需求侧响应中电价控制信号发出后的电价水平，另一方面改变 TCLs 集群中参与需求侧响应的个体数量。图 4-5 展示了在不同电价水平下，当参与需求侧响应的个体数量 N_d 分别设定为 30000 和 20000 时电力系统可靠性指标的变化。可以看出，随着电价水平的提高，LOLP 和 EPNS 都将降低，这是因为如第二章所示，在高水平电价下 TCLs 个体更倾向于降低自己的用电功率，从而为电力系统提供更多的备用容量，同时可以看出随着电价水平的提高，LOLP 和 EPNS 的下降速度逐渐降低，这是因为电价对 TCLs 个体的用电行为调整具有一定限度；同时可以看出在相同电价水平下，参与需求侧响应的个体数量的增大也会降低 LOLP 和 EPNS，这是因为参与需求侧响应的个体数量的增大直接导致了相同电价水平下 TCLs 集群的总备用容量的提高。因此，可以看出一定范围内的电价水平提高，以及参与需求侧响应的 TCLs 个体数量的提高，尽管对 TCLs 集群的影响方式不同，但是最终都提高了 TCLs 集群的备用容量，从而提高了电力系统的可靠性。这表明 TCLs 的需求侧响应是提高接入风电的电力系统保持供需平衡能力的有效手段。

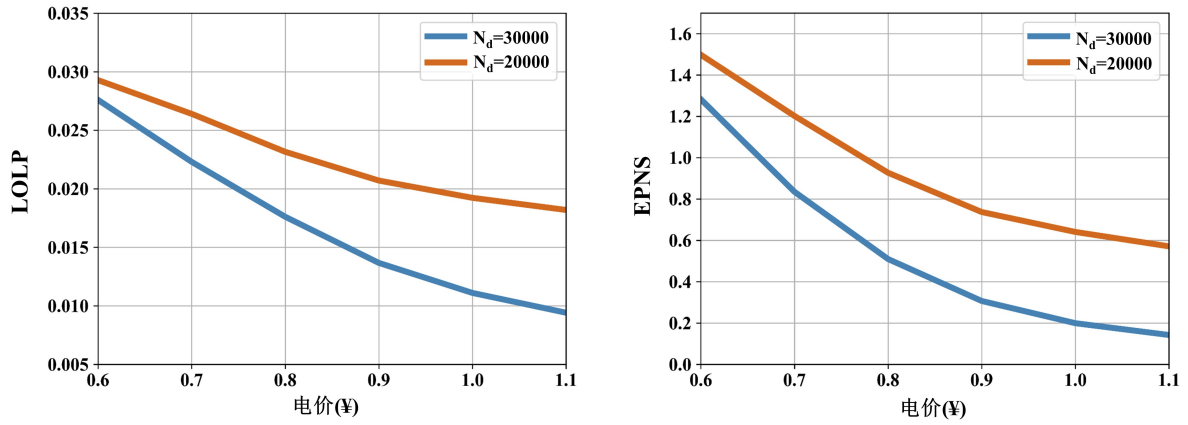


图 4-5 TCLs 对系统可靠性影响示意图

4.5. 本章小结

本章基于第二章对需求侧的 TCLs 参与需求侧响应以及第三章对供电侧的风电出力评估,综合考虑两者不确定性的影响对电力系统可靠性进行分析。具体地,在本章的可靠性分析中,通过聚合算法将 TCLs 集群等效为具有备用的发电机,通过 GANs 建立风电场出力与高斯随机变量的映射关系,通过伯努利随机变量描述电力系统元件运行状态,然后通过采用 CE 算法实现对随机变量的有效采样,实现对电力系统可靠性的高效分析评估。算例分析结果表明,本章所提的可靠性分析方法具有较高的计算效率,且 TCLs 参与需求侧响应和风电场装机容量对电力系统的可靠性具有显著影响, TCLs 的需求侧响应是提高接入风电的电力系统保持供需平衡能力的有效手段。

5. 总结与展望

5.1. 对本文的总结

随着全球气候恶化和污染问题的日益凸显,风力发电等清洁能源的发展和建设已经获得了世界性的高度重视。然而由于风电场出力具有随机性、间歇性的特点,风电的大量接入将对电力系统保持可靠稳定供电的能力提出巨大挑战。近年来以电力物联网技术的发展和电力市场化改革的推进作为契机,需求侧响应的发展也获得了很大的推动。以温控负荷为主要资源的需求侧响应,是提高接入风电的电力系统保可靠性的有效手段。然而温控负荷同样具有一定的响应程度不确定性。本文通过数据驱动算法,对需求侧的温控负荷不确定性和供电侧的风电场出力不确定性进行建模与评估,并以此为基础进行电力系统可靠性分析。全文总结如下:

(1) 提出了 TCLs 集群备用容量估计算法。首先对 TCLs 个体的运行模式和功率模型进行分析和构建,然后考虑到电网企业通常无法获取 TCLs 集群的所有响应特性和模型参数导致的响应信息不确定性,通过混合高斯模型和 K-means 聚类这两种数据驱动的聚合算法进行 TCLs 集群备用容量估计。算例分析表明,基于 K-means 聚类的聚合算法可以实现对 TCLs 集群备用容量的更加精确有效的估计。

(2) 提出了多风电场风速概率分布估计算法。在考虑风电出力波动性的前提下,基于 GANs 的对抗框架,博弈,通过历史风电数据对二元 ANN 进行博弈训练,在不依赖先验知识的条件下对风电场风速的概率分布进行估计,同时实现了同一区域处于不同位置的风电场风速之间的空间相关性估计。算例分析结果显示,本文提出的算法在单个风电场风速概率分布估计以及多风电场风速相关性估计这两方面都取得比常用的 Copula 方法更加准确的结果。

(3) 提出了考虑 TCLs 和风电不确定性的电力系统可靠性分析算法。基于对需求侧的 TCLs 参与需求侧响应以及供电侧的风电出力评估,综合考虑两者不确定性的影响对电力系统可靠性进行评估。通过数据驱动方法对电力系统可靠性分析评估问题进行重构以解决 CE 算法的使用障碍,然后通过系统状态采样、系

统负荷切除计算以及相关可靠性指标计算,实现了对考虑 TCLs 和风电不确定性的电力系统可靠性的高效评估。算例分析结果表明,本文所提的可靠性分析评估方法具有较高的计算效率,且 TCLs 参与需求侧响应和风电场装机容量对电力系统的可靠性具有显著影响,TCLs 的需求侧响应是提高接入风电的电力系统保持供需平衡能力的有效手段。

5.2. 下一步的研究展望

本文提出了数据驱动的考虑 TCLs 和风电不确定性的电力系统可靠性分析方法,基于本文的研究基础,可以从以下方面进行进一步的研究:

(1) 对于 TCLs 的运行模型,可以考虑采用更加精细化的模型,例如深入考虑室内外热交换的动态过程,以及通过实际工程试点采集相关的需求侧响应数据,对复杂的用户响应模式进行更加深入的研究,从而进一步提高对 TCLs 运行以及参与需求侧响应的建模和评估精度。

(2) 在建立电力系统可靠性分析问题,可以进一步考虑系统实时运行中的暂态过程对系统可靠性的影响,例如电力系统针对频率波动的调频过程、电力系统发电机组启停过程等因素对系统供电可靠性和安全运行的影响,从而进一步提高可靠性分析评估方法的准确程度。

参考文献

- [1] 汤涌. 电力系统安全稳定综合防御体系框架[J]. 电网技术, 2012, 36(8):1-5.
- [2] 张国华, 张建华, 彭谦, et al. 电网安全评价的指标体系与方法[J]. 电网技术, 2009, 33(8):30-34.
- [3] 程林, 何剑. 电力系统可靠性原理和应用[M]. 北京: 清华大学出版社, 2015.
- [4] Deng Y, Guo W. A Review of Investment, Financing and Policies Support Mechanisms for Renewable Energy Development[M]. Singapore: Springer, 2017.
- [5] Zhang D, Wang J, Lin Y, et al. Present situation and future prospect of renewable energy in China[J]. Renewable and Sustainable Energy Reviews, 2017, 76:865-871.
- [6] Global Wind Energy Council. Global wind report 2018[R/OL]. 2018. <https://gwec.net/wp-content/uploads/2019/04/GWEC-Global-Wind-Report-2018.pdf>
- [7] 罗磊. 抢占风电技术制高点, 服务全球能源互联网发展[J]. 中国电力企业管理, 2017(22):44-45.
- [8] 国家能源局. 国家能源局关于可再生能源发展“十三五”规划实施的指导意见[J]. 中国战略新兴产业, 2017(17):70-73.
- [9] B. Godfrey. Renewable electricity and the grid: the challenge of variability[M]. London: Routledge, 2009.
- [10] Liu Y, Du W, Xiao L, et al. Sizing a Hybrid Energy Storage System for Maintaining Power Balance of an Isolated System With High Penetration of Wind Generation[J]. IEEE Transactions on Power Systems, 2016, 31(4):3267-3275.
- [11] 丁一, 吴华华, 惠红勋, et al. 适合需求侧主动响应的中国电力市场机制的思考与建议[J]. 南方电网技术, 2016, 10(3):24-31.

- [12]Siano P, Sarno D. Assessing the benefits of residential demand response in a real time distribution energy market[J]. *Applied Energy*, 2016, 161:533-551.
- [13]中共中央国务院. 关于进一步深化电力体制改革的若干意见（中发〔2015〕9号）文[R], 2015.
- [14]Santamouris M, Cartalis C, Synnefa A, et al. On the impact of urban heat island and global warming on the power demand and electricity consumption of buildings—A review[J]. *Energy and Buildings*, 2015, 98:119-124.
- [15]Palensky P, Dietrich D. Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads[J]. *IEEE Transactions on Industrial Informatics*, 2011, 7(3):0-388.
- [16]Balijepalli, V M, Pradhan V, Khaparde S A, Shereef, R M (2011, December). Review of demand response under smart grid paradigm[C]. 2011 Innovative Smart Grid Technologies (ISGT), 2011.
- [17]Rahimi F, Ipakchi A. Demand Response as a Market Resource Under the Smart Grid Paradigm[J]. *IEEE Transactions on Smart Grid*, 2010, 1(1):82-88.
- [18]Torriti J, Hassan M G, Leach M. Demand response experience in Europe: Policies, programmes and implementation[J]. *Energy*, 2010, 35(4):1575-1583.
- [19]曹伟, 叶桂南, 周先哲, 宾灿, 陈根军. 新电改条件下电力市场需求侧响应交易机制分析[J]. *电网与清洁能源*, 2019, 35(10):50-55.
- [20]何德卫, 杨威, 陈皓勇, et al. 电力市场环境下需求侧响应相关问题的探讨[J]. *智慧电力*, 2018(6).
- [21]韩延民, 陈军, 肖明卫. 电力需求响应在部分试点城市实施情况调研分析[J]. *电力需求侧管理*, 2016(5):35-37.
- [22]R Sonderegger, Dynamic models of house heating based on equivalent thermal parameters[D]. Princeton University, 1978.

- [23] Wilson N W, Wagner B S, Colborne W G. Equivalent thermal parameters for an occupied gas-heated house[J]. ASHRAE Transactions, 91(CONF-850606-), 1985.
- [24] Zhang W, Lian J, Chang C Y, et al. Aggregated Modeling and Control of Air Conditioning Loads for Demand Response[J]. IEEE Transactions on Power Systems, 2013, 28(4):4655-4664.
- [25] Liu M, Shi Y. Model Predictive Control of Aggregated Heterogeneous Second-Order Thermostatically Controlled Loads for Ancillary Services[J]. IEEE Transactions on Power Systems, 2015, 31(3):1-9.
- [26] Behboodi S, Chassin D P, Djilali N, Crawford C. Transactive control of fast-acting demand response based on thermostatic loads in real-time retail electricity markets[J]. Applied Energy, 2018, 210:1310-1320.
- [27] Hanif S, Massier T, Gooi H B, et al. Cost Optimal Integration of Flexible Buildings in Congested Distribution Grids[J]. IEEE Transactions on Power Systems, 2016, 32(3):2254-2266.
- [28] Mathieu J L, Callaway D S. State Estimation and Control of Heterogeneous Thermostatically Controlled Loads for Load Following[C]. System Science (HICSS), 2012 45th Hawaii International Conference on. IEEE, 2012.
- [29] Koch S, Mathieu J L, Callaway D S. Modeling and control of aggregated heterogeneous thermostatically controlled loads for ancillary services[C]. 7th Proc. Power System Computation Conference, 2011.
- [30] Gu C, Yan X, Yan Z, et al. Dynamic pricing for responsive demand to increase distribution network efficiency[J]. Applied Energy, 2017, 205:236-243.
- [31] Lu N, Chassin D P. A state-queueing model of thermostatically controlled appliances[J]. IEEE Transactions on Power Systems Pwrs, 2004, 19(3):1666 - 1673.

- [32]Hui H, Ding Y, Liu W, Lin Y, Song Y. Operating reserve evaluation of aggregated air conditioners[J]. Applied energy. 2016, 196:218-228.
- [33]Xie D, Hui H, Ding Y, Lin Z. Operating reserve capacity evaluation of aggregated heterogeneous TCLs with price signals[J]. Applied energy, 2018, 216, 338-347.
- [34]Ding Y, Wang P, Goel L, et al. Long-Term Reserve Expansion of Power Systems With High Wind Power Penetration Using Universal Generating Function Methods[J]. IEEE Transactions on Power Systems, 2011, 26(2):0-774.
- [35]Sulaeman S, Benidris M, Mitra J, et al. A Wind Farm Reliability Model Considering Both Wind Variability and Turbine Forced Outages[J]. IEEE Transactions on Sustainable Energy, 2016:1-1.
- [36]Cheng L, Liu M, Sun Y, et al. A multi-state model for wind farms considering operational outage probability[J]. Journal of Modern Power Systems & Clean Energy, 2013, 1(2):177-185.
- [37]Dobakhshari A S, Fotuhi-Firuzabad M. A Reliability Model of Large Wind Farms for Power System Adequacy Studies[J]. IEEE Transactions on Energy Conversion, 2009, 24(3):792-801.
- [38]Billinton R, Allan R N. Reliability evaluation of engineering systems[M]. New York: Plenum Press, 1992.
- [39]Faulin J, Juan A A, Alsina S S M, Ramirez-Marquez J E. Simulation methods for reliability and availability of complex systems[M]. London: Springer Science & Business Media, 2010.
- [40]FranÇOis Vallee, Lobry J, Deblecker O. System Reliability Assessment Method for Wind Power Integration[J]. IEEE Transactions on Power Systems, 2008, 23(3):1288-1297.

- [41]Wen J, Zheng Y, Donghan F. A review on reliability assessment for wind power[J]. RENEWABLE & SUSTAINABLE ENERGY REVIEWS, 2009, 13(9):2485-2494.
- [42]Celik A N. Weibull representative compressed wind speed data for energy and performance calculations of wind energy systems[J]. Energy Conversion and Management, 2003, 44(19):3057-3072.
- [43]Ucar A, Balo F. Evaluation of wind energy potential and electricity generation at six locations in Turkey[J]. Applied Energy, 2009, 86(10):1864-1872.
- [44]Saeidi D, Mirhosseini M, Sedaghat A, et al. Feasibility study of wind energy potential in two provinces of Iran: North and South Khorasan[J]. Renewable & Sustainable Energy Reviews, 2011, 15(8):3558-3569.
- [45]Islam M R, Saidur R, Rahim N A. Assessment of wind energy potentiality at Kudat and Labuan, Malaysia using Weibull distribution function[J]. Energy, 2011, 36(2):985-992.
- [46]Guo Y, Gao H, Member, et al. A Combined Reliability Model of VSC-HVDC Connected Offshore Wind Farms Considering Wind Speed Correlation[J]. IEEE Transactions on Sustainable Energy, 2017, PP(99):1-1.
- [47]Zhou H, Wu H, Ye, C, et al. Integration Capability Evaluation of Wind and Photovoltaic Generation in Power Systems Based on Temporal and Spatial Correlations[J]. Energies, 2019, 12(1):171.
- [48]Buttler A, Dinkel F, Franz S, et al. Variability of wind and solar power – An assessment of the current situation in the European Union based on the year 2014[J]. Energy, 106:147-161.
- [49]Gao Y, Billinton R. Adequacy assessment of generating systems containing wind power considering wind speed correlation[J]. IET Renewable Power Generation, 2009, 3(2):217-226.

- [50]Kang J, Sun L, Sun H, et al. Risk assessment of floating offshore wind turbine based on correlation-FMEA[J]. Ocean Engineering, 2017, 129:382-388.
- [51]Xie Z Q, Ji T Y, Li M S, et al. Quasi-Monte Carlo Based Probabilistic Optimal Power Flow Considering the Correlation of Wind Speeds Using Copula Function[J]. IEEE Transactions on Power Systems, 2017:1-1.
- [52]Chen F, Li F, Wei Z, et al. Reliability models of wind farms considering wind speed correlation and WTG outage[J]. Electric Power Systems Research, 2015, 119:385-392.
- [53]Hagspiel S, Papaemannouil A, Schmid M, Andersson G. Copula-based modeling of stochastic wind power in Europe and implications for the Swiss power grid[J]. Applied energy, 2012, 96:33-44.
- [54]Billinton R. Power system reliability evaluation[M]. New York: Gordon and Breach, Science Publishers, Inc., 1970.
- [55]Roy Billinton. Bibliography on the Application of Probability Methods In Power System Reliability Evaluation[J]. IEEE Trans Power Syst, 1972, PAS-91(2):649-660.
- [56]Wang P, Billinton R. Time sequential distribution system reliability worth analysis considering time varying load and cost models[C]. IEEE, 1999:1046-1051.
- [57]Ding Y, Lisnianski A. Fuzzy universal generating functions for multi-state system reliability assessment[J]. Fuzzy Sets and Systems, 2008, 159(3):307-324.
- [58]Park J, Liang W, Choi J, et al. A probabilistic reliability evaluation of a power system including Solar/Photovoltaic cell generator[C]. Power & Energy Society General Meeting, 2009. PES '09. IEEE. IEEE, 2009.
- [59]Zhang Y, Wang L, Xiang Y, et al. Power System Reliability Evaluation With SCADA Cybersecurity Considerations[J]. IEEE Transactions on Smart Grid, 2015, 6(4):1707-1721.

- [60] 杨森, 杜文娟, 王旭斌, et al. 风火-需求侧响应协调频率控制方法[J]. 电网技术, 2017(03):173-181.
- [61] Moshari A, Ebrahimi A, Fotuhi-Firuzabad M. Short-Term Impacts of DR Programs on Reliability of Wind Integrated Power Systems Considering Demand-Side Uncertainties[J]. IEEE Transactions on Power Systems, 2015, 31(3):1-10.
- [62] He, M, Murugesan S, Zhang, J. A multi-timescale scheduling approach for stochastic reliability in smart grids with wind generation and opportunistic demand[J]. IEEE Transactions on Smart Grid, 2013, 4(1):521-529.
- [63] Silva A M L D, Gonzalez-Fernandez R A, Flavio S A, et al. Composite reliability evaluation with renewable sources based on quasi-sequential Monte Carlo and cross entropy methods[C]. 2014 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS). IEEE, 2014.
- [64] Singh C, Mitra J. Composite system reliability evaluation using state space pruning[J]. IEEE Transactions on Power Systems, 1997, 12(1):471-479.
- [65] Borges C L T, Falcao D M, Mello J C O, et al. Composite reliability evaluation by sequential Monte Carlo simulation on parallel and distributed processing environments[J]. IEEE Transactions on Power Systems, 2001, 16(2):203-209.
- [66] Rubinstein R Y, Kroese D P. The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation and machine learning[M]. New York: Springer Science & Business Media, 2013.
- [67] Martins L D S A, Costa Castro J F, Gonzalez-Fernandez R A. Spinning Reserve Assessment Under Transmission Constraints Based on Cross-Entropy Method[J]. IEEE Transactions on Power Systems, 2015:1-9.
- [68] Wang Y, Guo C, Wu Q H. A cross-entropy-based three-stage sequential importance sampling for composite power system short-term reliability evaluation[J]. IEEE Transactions on Power Systems, 2013, 28(4):4254-4263.

- [69]Tómasson E, Söder L. Improved importance sampling for reliability evaluation of composite power systems[J]. IEEE Transactions on Power Systems, 2017, 32(3):2426–2434.
- [70]Wang Y, Guo C, Wu Q, et al. Adaptive sequential importance sampling technique for short-term composite power system adequacy evaluation[J]. IET Generation Transmission & Distribution, 2014, 8(4):730-741.
- [71]Roussa G G. An introduction to probability and statistical inference[M]. Orlando: Elsevier, 2003.
- [72]Kutz J N, Brunton S L, Brunton B W, Proctor J L. Dynamic mode decomposition: data-driven modeling of complex systems[M]. Philadelphia: Society for Industrial and Applied Mathematics, 2016.
- [73]Kubat, M. An introduction to machine learning (Vol. 2)[M]. Cham: Springer International Publishing, 2017.
- [74]Stamp, M. Introduction to machine learning with applications in information security[M]. Boca Raton : Chapman and Hall/CRC, 2017.
- [75]I Goodfellow, Bengio Y, Courville A, Deep learning[M]. London: MIT Press, 2016.
- [76]Wang Y, Liu M, Bao Z. Deep learning neural network for power system fault diagnosis[C]. 2016 35th Chinese Control Conference (CCC). IEEE, 2016.
- [77]曹靖,陈陆燊,邱剑,王慧芳,应高亮,张波.基于语义框架的电网缺陷文本挖掘技术及其应用[J].电网技术,2017,41(02):637-643.
- [78]Wang Q, Li F, Tang Y, Xu, Y. Integrating Model-driven and Data-driven Methods for Power System Frequency Stability Assessment and Control[J]. IEEE Transactions on Power Systems, 2019, 34(6): 4557-4568.
- [79]Chae Y T, Horesh R, Hwang Y, et al. Artificial Neural Network Model for Forecasting Sub-hourly Electricity Usage in Commercial Buildings[J]. Energy and Buildings, 2015, 111:184-194.

- [80]Ruelens F, Claessens B, Vandael S, et al. Residential demand response of thermostatically controlled loads using batch Reinforcement Learning[J]. IEEE Transactions on Smart Grid, 2017, 8(5):2149-2159.
- [81]Isaac M, Vuuren D P V. Modeling global residential sector energy demand for heating and air conditioning in the context of climate change[J]. Energy Policy, 2009, 37(2):507-521.
- [82]Hongxun H, Yi D, Menglian Z. Equivalent Modeling of Inverter Air Conditioners for Providing Frequency Regulation Service[J]. IEEE Transactions on Industrial Electronics, 2018:1-1.
- [83]Yi D, Yonghua S, Hongxun H, et al. Integration of Air Conditioning and Heating into Modern Power Systems[J]. Singapore: Springer, 2019.
- [84]Xuan G, Zhang W, Chai P. EM algorithms of Gaussian mixture model and hidden Markov model[C]. Proceedings 2001 International Conference on Image Processing (ICIP 2001), 2001.
- [85]Wong J, A Hartiganm A. Algorithm AS 136: A K-Means Clustering Algorithm[J]. Journal of the Royal Statistical Society. Series C (Applied Statistics), 1979, 28(1):100-108.
- [86]Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]. 2014 Neural information processing systems, 2014.
- [87]Hinrich S, Christopher D M. Foundations of Statistical Natural Language Processing[M]. Cambridge, Mass: MIT Press, 1999.
- [88]Montavon G, Orr G, Müller K R. Neural networks: tricks of the trade[M]. New York: Springer, 2012.
- [89]Durante F, Sempi C. "Copula theory: an introduction," In Copula theory and its applications[M]. Berlin: Springer, 2010.
- [90]C Spearman. The Proof and Measurement of Association between Two Things[J]. The American Journal of Psychology, 1987.

- [91]Rubinstein R Y, Kroese D P. The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation and machine learning[M]. New York: Springer Science & Business Media, 2013
- [92]Singh C, Jirutitijaroen P, Mitra J. Electric power grid reliability evaluation: models and methods[M]. New Jersey:Wiley-IEEE Press, 2018.
- [93]Diaf S, Notton G, Belhamel M, et al. Design and techno-economical optimization for hybrid PV/wind system under various meteorological conditions[J]. Applied Energy, 2008, 85(10): 968-987.
- [94]Reliability Test System Task Force of the Application of Probability Methods Subcommittee. IEEE reliability test system[J]. Power Apparatus & Systems IEEE Transactions on, 1979, 98(6):2047-2054.
- [95]Jia H, Ding Y, Song Y, et al. Operating reliability evaluation of power systems considering flexible reserve provider in demand Side. IEEE Transactions on Smart Grid, 2019, 10(3):3452–2464.

作者简介

庄欣然，男，汉族，1994年12月生，研究方向为智能电网与电力系统可靠性分析。

2013.9-2017.6 浙江大学电气工程学院，电气工程及其自动化，本科；

2017.9-2020.3 浙江大学电气工程学院，电力系统及其自动化，硕士。

攻读硕士学位期间的学术成果

1. 学术论文

[1] **Xinran Z**, Chengjin Y, Yi D, et al. Data-driven Efficient Reliability Evaluation of Power Systems with Wind Penetration: An Integrated GANs and CE Method[J]. IET Generation, Transmission & Distribution, 2019. (SCI期刊, 已录用)

[2] **Xinran Z**, Chengjin Y, Yi D, et al. Data-Driven Reserve Allocation With Frequency Security Constraint Considering Inverter Air Conditioners[J]. IEEE Access, 7:120014-120022. (SCI期刊, 已发表)

[3] **Xinran Z**, Chengjin Y, Yi D, et al. Reliability Evaluation of Power Systems Considering Integration of Renewable Energy and Aggregated TCLs[C]. 2019 4th International Conference on Power and Renewable Energy (ICPRE), 2019. (EI会议, 已录用)

2. 项目工作

[1] 国家重点研发计划项目，电力需求侧主动响应的市场机制研究（2016YFB090103），2017年9月至2018年12月，项目核心成员。

[2] 电力市场环境下供需双侧互动的能源生态智能交互系统研究，杭州，浙江，2017年12月至2019年6月，项目核心成员。

[3] 电网新能源承载力分析和优化调度关键技术研究应用，宁波，浙江，2018年6月至2019年11月，项目核心成员。

- [4] 多能互补协调优化规划设计关键技术研究，南京，江苏，2018 年 11 月至 2019 年 10 月，项目核心成员。
- [5] 基于物联网的城市配电网节能规划与集群控制技术研究，深圳，广东，2019 年 4 月至今，项目核心成员。