



数字时代图书馆学情报学青年论丛

(第二辑)

基于计算智能技术的 电力负荷预测理论及应用

Computational Intelligence-based Electricity Load Forecasting:
Theory and Applications

胡忠义 著



WUHAN UNIVERSITY PRESS

武汉大学出版社



数字时代图书馆学情报学青年论丛

(第二辑)

本书的主要研究工作获得国家自然科学基金“基于多输出支持向量回归的预测技术研究”(71571080)、“基于集成学习的区间型电力负荷预测技术研究”(71601147)以及中国博士后科学基金“基于多输出支持向量机的区间型电力负荷预测研究”(2015M582280)等基金项目的资助

本书的出版得到国家自然科学基金项目“基于集成学习的区间型电力负荷预测技术研究”(71601147)的资助

基于计算智能技术的 电力负荷预测理论及应用

Computational Intelligence-based Electricity Load Forecasting:
Theory and Applications

胡忠义 著



WUHAN UNIVERSITY PRESS

武汉大学出版社

图书在版编目(CIP)数据

基于计算智能技术的电力负荷预测理论及应用/胡忠义著. —武汉: 武汉大学出版社, 2019.1

数字时代图书馆学情报学青年论丛. 第二辑

ISBN 978-7-307-20613-7

I.基… II.胡… III.电力系统—负荷(电)—预测—研究
IV.TM715

中国版本图书馆 CIP 数据核字(2018)第 254943 号

责任编辑:陈 豪

责任校对:李孟潇

版式设计:马 佳

出版发行: 武汉大学出版社 (430072 武昌)

(电子邮件: cbs22@whu.edu.cn 网址: www.whu.com.cn)

印刷: 北京虎彩文化传播有限公司

开本: 720×1000 1/16 印张: 13.75 字数: 198 千字 插页: 1

版次: 2019 年 1 月第 1 版 2019 年 1 月第 1 次印刷

ISBN 978-7-307-20613-7 定价: 48.00 元



版权所有, 不得翻印; 凡购我社的图书, 如有质量问题, 请与当地图书销售部门联系调换。

前 言

随着电力系统运营环境的日趋复杂,中国经济新常态下电力体制改革的不深化和竞争性电力交易市场的逐步构建与完善,电力负荷的准确预测在提升电力系统运营管理水平和提高电力企业经济效益等方面扮演着特别重要的角色,并已成为管理科学在电力系统运营与管理中的重要研究课题。

尽管以往研究针对各种负荷预测问题提出了许多优秀的负荷预测理论和方法,但无论采用哪种预测方法,预测模型的性能都受到多种关键技术细节(如模型的输入特征选择、参数选择、模型选择等)的影响。本书综合运用多种计算智能技术,提出了基于计算智能的电力负荷预测理论框架(CILF 框架);并在此框架下围绕建模过程中的多个关键技术问题开展了一系列研究;同时,结合电力负荷预测研究中短期负荷预测、区间型负荷预测等典型预测问题进行了应用研究。此外,结合本书基于计算智能的负荷预测理论框架,特别是在特征选择、参数优化、模型选择等具体技术问题上的研究成果,设计了负荷预测支持系统,为决策者提供决策工具。综观全书,主要的研究工作和创新性成果如下:

①通过综合使用多种计算智能技术,提出基于计算智能的电力负荷预测理论框架。该框架采用计算智能等领域先进的建模技术(如神经网络、支持向量机等)构建负荷预测模型,融合计算智能领域先进的进化计算等优化算法,自适应地优化负荷预测模型的输入特征、参数等关键技术细节,以提升负荷预测模型的准确性与自

适应性。

②针对基于支持向量机(support vector regression, SVR)的电力负荷预测模型中参数优化的问题,提出了基于萤火虫算法的文化基因算法,并将其用来寻优 SVR 负荷预测模型的参数。基于两个实际电力市场的负荷预测案例表明,该方法能自适应地寻优预测模型的参数,并有效提高模型的预测精度。

③围绕负荷预测模型中的特征选择问题,为简化 Wrapper 方法的搜索空间、整合 Filter 方法的优势,设计了 Filter-Wrapper 混合输入选择方法,并结合短期电力负荷预测问题进行了实证研究。结果表明,该混合特征选择算法能高效地筛选出较少的特征子集,并能有效提高 SVR 负荷预测模型的预测精度。

④针对已有的负荷预测研究更多关注单值预测的情形,提出了基于多输出支持向量机(MSVR)的区间型负荷预测模型,并针对区间型负荷预测模型中的特征选择问题,基于文化基因算法(memetic algorithms, MA)设计了适应区间型负荷数据特征的特征选择算法。结合实际负荷预测案例的实证研究表明,提出的 MSVR-MA 模型不仅能自适应地筛选输入特征,而且得到的预测精度明显优于其他对比模型。

⑤针对以往研究一般将特征选择和参数优化分别采用不同的方法进行优化的问题,本书提出了统一框架下的广义模型选择问题(generalized model selection problem)。围绕模型选择这种混合连续-离散优化问题,进一步设计了基于 CLPSO 的混合变量 MA 算法,并用来优化提出的模型选择问题。结合电力负荷预测这一电力市场中的传统热点问题,通过对比多种不同的模型选择策略,统一框架下的广义模型选择的优势及提出的 CLPSO-MA 算法在解决模型选择问题方面的优势均得到了验证。

⑥基于本书在预测模型参数优化、特征选择、模型选择等方面的研究成果,通过封装 CILF 框架,设计了基于计算智能的电力负荷预测支持系统。该系统整合了丰富的优秀定量预测模型和专业的专家经验知识,集成先进的模型参数优化、特征选择和模型选择等模型构建中的关键问题解决方案,通过友好的界面和用户交互使得

用户能够较好地解决负荷预测这类复杂问题。

本书的研究工作得以顺利完成，要特别感谢华中科技大学张金隆教授和鲍玉昆教授、澳大利亚 The University of Newcastle 的 Raymond Chiong 高级讲师、华中农业大学熊涛副教授等的指导与合作；感谢武汉大学信息管理学院各位领导和老师在工作上的帮助与支持；感谢武汉大学出版社和詹蜜编辑对本书的出版给予的极大协助与支持。本书的主要研究工作是在国家自然科学基金(71571080、71601147)、中国博士后科学基金(2015M582280)等基金项目的资助下完成的。全书的出版得到国家自然科学基金项目(71601147)的资助。

限于作者能力和水平，书中难免存在许多不足或有待进一步深入研究的地方，敬请各位专家、读者批评指正。

胡忠义

2018 年 2 月

目 录

第 1 章 绪论	1
1.1 研究背景、目的与意义	1
1.2 研究现状	4
1.3 本书结构与主要内容	21
第 2 章 基于计算智能的电力负荷预测理论框架与方法	25
2.1 引言	25
2.2 基于计算智能的电力负荷预测理论框架(CILF)	26
2.3 人工神经网络	29
2.4 支持向量机	32
2.5 粒子群优化算法	43
2.6 文化基因算法	44
2.7 本章小结	46
第 3 章 基于 MA 的电力负荷预测模型参数优化	47
3.1 引言	47
3.2 支持向量机参数简述	50
3.3 基于 MA 算法的 SVR 参数优化	51
3.4 实验设置	56
3.5 实验结果与分析	58
3.6 本章小结	64

第 4 章 基于混合特征选择技术的负荷预测模型	66
4.1 引言	66
4.2 混合 Filter-Wrapper 特征选择算法的构建过程	70
4.3 实验设计	75
4.4 实验结果与分析	84
4.5 本章小结	95
第 5 章 基于 MA 算法的特征选择技术及区间型负荷预测 模型	96
5.1 引言	96
5.2 问题描述与分析	99
5.3 MSVR-MA 预测模型的构建	104
5.4 实验设计	111
5.5 实验结果与分析	115
5.6 本章小结	129
第 6 章 基于混合变量 MA 的电力负荷预测模型选择研究 ...	131
6.1 引言	131
6.2 问题描述	134
6.3 基于 MA 算法的广义模型选择	136
6.4 实验设计	143
6.5 实验结果与分析	146
6.6 本章小结	155
第 7 章 电力负荷预测支持系统	156
7.1 引言	156
7.2 系统需求分析	158
7.3 系统结构设计	159
7.4 系统逻辑流程	165
7.5 本章小结	166

第 8 章 总结与展望	168
8.1 全书总结	168
8.2 研究展望	170
参考文献	172
附录 A 区间数运算法则	202
附录 B 区间数的描述性统计	204
附录 C 区间多层感知器	206
附录 D 指数平滑与区间型 Holt 模型	209

第1章 绪 论



1.1 研究背景、目的与意义

1.1.1 研究背景

电力系统负荷预测通过对历史负荷以及经济、气象等影响因素的分析,探讨负荷的发展变化规律及其与影响因素间的关系,并结合经济、气象等影响因素的未来发展趋势,对电力需求做出尽可能准确的预先估计和推测^[1]。

科学的预测是正确决策的依据和保障。随着电力系统运营环境的日趋复杂和电力体制改革的不断深化,系统负荷的准确预测不仅有助于提升电力系统的运营管理水平,还有助于提高电力企业的经济效益。首先,电力负荷预测在电力系统规划和运营中起着至关重要的作用。它是电力企业进行基础设施建设规划、电力生产与输送计划、燃料购买与运输计划、财务收支计划等日常经营规划过程中的重要参考信息。其次,负荷预测还是保证电力系统可靠供电和经济运行的基础。研究表明,预测误差增加1%会导致每年1000万欧元运行成本的增加^[2]。在中国经济“稳增长、调结构”的新常态下,科学准确的负荷预测有助于我国节约能源消耗和节能减排任务的完

成,有助于促进电力企业高效运行,优化电力工业布局,以满足国民经济健康持续发展和居民生活水平不断提高的需要^[3]。再次,随着中国电力体制改革的深入和电力市场化的不断加强,尤其是随着国家《能源发展战略行动计划(2014—2020年)》和《关于进一步深化电力体制改革的若干意见》的颁布和实施,我国竞争性电力交易市场将逐步构建和完善,电力负荷的预测将在电力交易、市场营销等部门得到广泛重视,将成为电力企业售电、购电的重要参考信息,是供电企业进行调峰辅助服务的基础信息,是电网安全、经济、高效运行的重要保障。因此,提高负荷预测的精度是电力市场化运行后的必然要求。总之,负荷预测不仅涉及电力系统的规划与设计,还涉及电力系统经济性与安全性运营、电力市场交易等多个方面,是现代电力系统运营和管理中一个非常重要的研究领域^[3]。

目前国内外学者不遗余力地针对负荷预测研究,相继提出各种负荷预测理论与方法。以统计学为学科基础的传统预测模型,如自回归移动平均模型^[4]、指数平滑法^[5,6]和其他回归分析模型^[7-9],因其本质是统计数据规律服从线性特征,而在非线性、非平稳的负荷预测应用中不能达到人们对预测结果的预期。基于此,面向非线性数据预测的方法与技术成为负荷预测研究的热点,其中基于数据驱动的机器学习技术(如人工神经网络、支持向量机等)在负荷预测领域备受关注^[2,10]。此外,混合模型通过组合不同预测模型能有效弥补单个模型在预测精度上的缺陷,在负荷预测领域也得到了广泛研究^[11,12]。

尽管各种优秀的预测技术不断涌现,很多研究同时表明,无论是传统统计学的时间序列预测方法,还是基于数据驱动的机器学习方法和混合模型,其预测性能都受多种技术细节的影响(如模型的输入特征、模型结构与参数选择、模型选择等)^[2,10]。这些理论上很优秀的学习算法在实际预测中因为不恰当的模型设置而使得性能差强人意^[2,10],即使是一些商业的预测软件也会因为不同的模型设置而得出不同的结果^[13]。因此,通过解决模型的输入选择、参数选择和模型选择问题来提高负荷预测模型的预测性能,成为发展负

荷预测理论和方法的必经之路。

计算智能(computational intelligence, CI)技术以数据和模型为基础、以计算为手段建立数据驱动的智能模型,进而到达求解问题、实现对智能的模拟和认识的目的。它由若干相对独立的分支领域综合发展而来,目前已经形成的具有代表性的分支有神经网络、模糊系统、进化计算等领域。基于一种或多种计算智能技术构建的智能模型往往具有自学习、自组织、自适应的特征,以及简单、通用、鲁棒性强、适于并行处理的优点。计算智能技术已经在模式识别、智能控制、经济预测、数据挖掘、医疗诊断等领域得到广泛应用。

1.1.2 研究目的

本书的研究目的是针对电力负荷预测这一电力系统运营管理中的重要研究问题,结合计算智能等领域较为先进的预测建模技术(如支持向量机、神经网络等)和进化计算技术(如文化基因算法、粒子群优化算法、萤火虫算法等),设计基于计算智能技术的负荷预测理论框架(computational intelligence-based load forecasting, CILF);通过解决CILF框架中的关键技术问题(如输入选择、模型参数优化、模型选择等问题),构建智能化的自适应负荷预测模型,以提升负荷预测模型的性能。为有效解决这些问题,总结电力负荷预测及建模构建的研究进展和应用领域,深入研究模型构建和优化的流程与机理,设计用于解决负荷预测模型构建中的输入选择、模型参数选择、模型选择等问题的智能算法。同时,结合电力负荷预测研究中的短期负荷预测、区间型负荷预测等典型预测问题开展应用研究。最后,结合以上研究成果,设计基于计算智能技术的电力负荷预测支持系统,通过良好的系统结构设计 with 自适应学习机制设计,实现信息咨询、数据分析与处理、负荷预测等功能,为决策者提供决策辅助工具。



1.1.3 研究意义

本研究提出基于计算智能的负荷预测理论框架,采用先进的计算智能理论与方法建立智能的负荷预测模型,并构建智能负荷预测支持系统,以提升负荷预测的准确性与智能性,具有较好的理论贡献,并为电力系统安全高效的运营及运营企业经济效益的提高提供有效的决策模型与决策支持。

同时,本研究在 CILF 框架下,针对电力负荷预测建模中的参数优化、特征选择、模型选择等关键技术问题,分别构建为连续变量优化、0/1 离散优化、混合连续-0/1 离散变量优化等优化问题,通过引入先进的进化计算技术,尝试设计与负荷预测模型中不同类型的优化问题相适用的优化算法,并结合所选的典型负荷预测问题,对所提出的优化技术及基于优化技术的预测模型进行验证。从优化算法的角度,本书涵盖多种针对不同优化问题(连续变量优化、0/1 离散优化、混合连续-0/1 离散变量优化等)的进化算法的设计,既在一定程度上推动了计算智能领域里进化计算理论与方法的发展,又充分挖掘其在优化负荷预测模型与提升负荷预测模型性能中的潜力。

总之,通过本书的研究,可以有效克服以往负荷预测模型中预测性能差强人意的弱势,对于发展完善进化算法和具有自适应能力的负荷预测模型及预测支持系统具有创新性基础研究的理论意义,同时也表现出良好的应用研究的现实意义。



1.2 研究现状

1.2.1 电力负荷预测研究概况

按照负荷预测时间长短(time horizon or the load time),可将负

负荷预测问题分为超短期、短期、中期和长期四类^[14, 15]。四种负荷预测问题因提供的负荷预测的提前期不同而对电力系统的运行与规划具有不同方面的重要意义^[16]。超短期负荷预测(very short-term load forecasting, VSTLF)主要是预测未来几分钟至几小时内的负荷值,主要用于电网的安全监视和电力系统的潮流控制等紧急情况处理。短期负荷预测(short-term load forecasting, STLF)主要用于预测未来几小时、一天至几天的电力负荷,主要用于指导机组调度与最优组合、经济调度、电力市场运营及交易等。短期负荷预测是负荷预测中关注最广泛的预测问题。中期负荷预测(mid-term load forecasting, MTLF)旨在预测提前一个月至一年的负荷需求,主要用于指导电力系统的中期运行计划,如安排月度检修计划与运行方式、水库调度计划、电煤计划等规划工作,以降低运行成本和提高供电可靠性。长期负荷预测(long-term load forecasting, LTLF)是预测一年以上的负荷需求,主要用于制定电力系统发展计划,指导电源和电网的建设规划,以满足国民经济增长和人民生活水平提高对负荷的需求。

鉴于电力负荷预测的重要性,各种负荷预测理论和方法层出不穷,大致可分为两类:传统预测方法和基于人工智能方法的现代预测技术^[14, 15, 17, 18]。

传统的预测方法主要包括以统计学为基础的回归模型和时间序列预测模型。回归模型通过构建负荷与相关影响因素的经验回归方程,研究各变量之间的相互关系,并对历史负荷进行拟合和预测^[7, 9, 19]。时间序列预测方法将负荷数据看做一个按季节、周、天以及小时周期变化的时间序列,建立描述负荷变化规律的数学模型。时间序列预测模型在负荷预测研究中获得了广泛应用^[20-25]。如唐俊杰等人采用 ARMA(auto-regressive moving average model)负荷预测模型实现了从自动定阶到节假日负荷的预测^[26];叶瑰昀等人构建了基于加权最小二乘法的 ARMA 日负荷预测模型^[27];Pappas 等人使用 ARMA 构建的负荷预测模型在预测中期负荷和提前一周的负荷上取得了较好的性能^[22];Huang 等人构建了基于粒子群优化算法的 ARMAX(auto-regressive and moving average with exogenous

variables)短期负荷预测模型,该模型使用粒子群优化算法优化模型的阶数和系数^[23];而Wang等人基于ARMAX构建了短期负荷预测模型,并使用混合遗传-粒子群算法优化模型参数,获得了比基于传统梯度下降法的预测模型更优的预测性能^[24];为了提高时间序列负荷预测模型的精度,一些研究尝试将时间序列负荷预测模型和先进的人工智能预测方法如模糊逻辑、人工神经网络等结合起来,并取得了很大的成功^[20,28]。

传统的预测方法因其本质是统计数据规律服从线性特征,而在非线性非平稳的负荷预测应用中不能达到人们的预期。基于此,面向非线性数据预测的方法与技术成为负荷预测研究的热点,其中基于人工智能的方法在负荷预测领域备受关注,如模糊理论^[29-31]、专家系统^[32-34]、小波分析^[35,36]、神经网络^[37,38]、支持向量机^[39-41]和混合模型^[11,12]等。其中,神经网络和支持向量机在负荷预测研究中得到了最广泛的应用。

神经网络(neural networks, NN)具有较好的自学习能力和复杂的非线性数据拟合能力,能自适应地捕捉大量非结构性、非精确性的变化规律,在电力负荷预测中得到广泛关注^[2]。Mandal等人提出的基于NN和相似日的负荷预测模型,可以有效地预测未来1~6小时的负荷^[42]。运用粒子群算法自适应地调整网络权重,Bashir和El-Hawary提出了自适应的人工神经网络预测模型,结果表明该模型比传统的BP神经网络能更有效地提高负荷预测的精度^[43]。Lauret等人运用贝叶斯神经网络模型预测短期负荷,并取得了比传统NN更好的预测精度^[44]。利用改进的和音算法(harmony search)作为学习算法,神经网络取得了比多个基准模型更好的负荷预测精度^[45]。Hippert等人对神经网络在短期负荷预测的应用作了较全面的综述与评估^[2]。然而,NN预测模型的模型结构非常复杂,许多参数(如网络层数和神经元个数)均需要依赖经验设置,已有的学习算法极易收敛到局部最优,因而限制了神经网络模型的应用^[2,46]。

支持向量机(support vector machines, SVMs)是Vapnik等人在统计学习理论的基础上提出的一种新型的机器学习技术^[47]。不同

于神经网络算法,通过使用结构风险最小化原则代替传统的经验最小化原则,使得 SVMs 能较好地解决有限样本的学习问题和过学习问题;核函数的使用使得 SVMs 把非线性问题转化为线性问题来解决,有效降低了算法的复杂度。与 NN 相比, SVMs 具有收敛速度快、预测能力强、有效避免过学习和陷入局部极小等问题的优点,在负荷预测中取得了广泛的关注^[10]。例如,在 EUNITE 中期负荷预测竞赛中, Chen 等人使用支持向量机构建的负荷预测模型取得了第一名的成绩^[39]。Fan 和 Chen 构建了基于自组织神经网络-支持向量机的短期负荷预测模型,该模型首先使用自组织神经网络选择训练数据,然后分别构建各小时的支持向量机负荷预测模型,取得了比单独的支持向量机负荷预测更优的精度^[48]。Elattar 等人提出使用局部加权 SVR 模型构建短期负荷预测模型,该模型比文献中已发表的模型有更优的预测能力^[40]。Wang 等人提出的 TF- ε -SVR 模型,通过使用趋势修正和季节调整以提高基于 SVR 的负荷预测模型的精度^[49]。电力负荷预测已经是支持向量机及其变种(如最小二乘支持向量机)在回归预测中最广泛的应用领域之一^[10]。

尽管各种优秀的预测技术不断涌现,很多研究同时表明,无论是传统统计学的时间序列预测方法,还是基于数据的机器学习方法和混合模型,其预测性能都受多种技术因素的影响(如模型的输入特征、模型结构与参数选择、模型选择等)^[2,10]。相关研究表明,一些理论上很优秀的学习算法在实际预测中因为不恰当的模型设置而表现差强人意,即使是一些商业的预测软件也会因为不同的模型设置而得出不同的结果。因此,通过解决模型的输入选择、参数选择和模型选择问题来提高负荷预测模型的预测性能,成为发展负荷预测理论和方法的必经之路。

此外,值得指出的是,以往关于负荷预测的研究大多是基于单值预测的确定性预测,即仅给出某一预测时间点的确定性预测值。然而,电力系统中蕴含着各种不确定性因素,决策工作因而面临各种不确定性的风险。为了方便决策人员对电力系统进行风险分析和可靠性评估,一些研究关注电力负荷的区间预测(prediction interval)^[50-53]、概率预测^[54,55]和区间型预测^[56-58]。在这些预测负荷

不确定性的研究中,区间型负荷预测是最近几年新兴的一种方法。它研究区间型负荷的预测方法,该方法在训练数据中以区间型数据的方式直接考虑负荷的不确定性和波动性,与传统的基于单值负荷的建模方式相比可以有效降低数据的随机波动性^[58],因此受到一些研究者的关注^[56]。正如 García-Ascanio 和 Maté 提到的,区间型负荷预测为电力企业制定电力系统规划和运作决策提供了一个非常重要的风险管理工具^[56]。此外,预测未来的区间负荷还是电力系统进行负载潮流分析、输电网扩展规划需要参考的重要信息。鉴于区间型负荷预测在实践中的重要作用,本书在后面章节除了关注典型的负荷预测尤其是短期负荷预测的建模研究外,也探索对区间型负荷预测这一新兴的负荷预测领域的研究。

1.2.2 支持向量机参数优化研究概况

支持向量机具有坚实的理论基础,在预测研究中得到了越来越广泛的关注。其中,电力负荷预测是支持向量机回归预测中最广泛的应用领域之一^[10],许多研究基于 SVM 开发的负荷预测模型比传统的时间序列预测方法、神经网络等模型具有更优的精度,因此,本书重点关注基于支持向量机的负荷预测模型的研究。

在使用 SVM 构建预测模型时,研究中普遍关注的问题是核函数和模型参数的选择。考虑到径向基(RBF)核函数优越的性能、较少的参数和简单的计算复杂度,以往研究一般首选 RBF 核函数^[59,60]。当采用 RBF 核函数时,SVM 预测模型里共有三个参数需要调整。其中,惩罚系数 C 控制经验误差与模型复杂度(VC 维)的平衡。 δ 代表支持向量对输入变量变化的敏感程度,且其影响是非单调的。过大的 δ 会使支持向量反应迟钝,不能随着输入的变化而迅速调整;反之,过小的 δ 则对输入过于敏感,以致干扰对支持向量的影响,甚至导致对样本学习的失败。 ε 是不敏感损失函数中拟合管道的宽度,代表误差的边界。 ε 值越大,导致支持向量的个数越少,进而精度越低^[60]。因而,为保证负荷预测模型的精度和泛化性能,对惩罚因子 C 、核参数 δ 和不损失函数带宽 ε 三个参数的