

电力负荷预测技术 及其应用

牛东晓 曹树华 赵磊 张文文 编著

中国电力出版社

前 言

电力负荷预测是电力系统调度、用电、计划、规划等管理部门的重要工作之一。提高负荷预测技术水平,有利于计划用电管理,有利于合理安排电网运行方式和机组检修计划,有利于节煤、节油和降低发电成本,有利于制定合理的电源建设规划,有利于提高电力系统的经济效益和社会效益。因此,负荷预测已成为实现电力系统管理现代化的重要内容之一。

负荷预测的核心问题是预测的技术方法,或者说是预测的数学模型。随着现代科学技术的快速发展,负荷预测技术的研究也在不断深化,各种各样的负荷预测方法不断涌现,从经典的单耗法、弹性系数法、统计分析法,到目前的灰色预测法、专家系统法和模糊数学法,甚至到刚刚兴起的神经网络法、优选组合法和小波分析法,它们都有各自的研究特点和适用条件,不弄清其模型结构和适用范围而盲目地生搬硬套,可能不会取得理想的预测效果,反而事倍功半。因此,我们希望通过编写此书,帮助大家掌握各类负荷预测方法,研究其模型结构、功能特点、适用范围,了解各类方法在电力系统的实际应用情况,并且介绍几类处于科研前沿的负荷预测技术,以利于搞好负荷预测的工作,提高负荷预测人员的技术水平。

本书是应电力工业部全国计划用电办公室的要求,结合我们多年的科研及教学工作体会编写而成的。编写过程中充分考虑了我国电力系统负荷预测工作的需要,可供电力类、技术经济类、计划统计类、管理工程类专业本科生和研究生使用,也可供电力系统的调度、用电、计划、规划和研究等管理部门的科技人员参考使用。本教材之外我们还编有各类负荷预测软件,有《电网日负荷预测系统》、《重大节日负荷预测系统》、《用电负荷预测系统》、《行业电量预测系统》、《周、月、季、年负荷预测系统》等,都已在华北电管局和河北省电力局等部门实际应用。

在本书的编写过程中,始终受到了电力工业部计划用电办公室李阳主任的关怀和指导,受到了华北电管局用电营业部师坦丁、于淑云两位领导的支持和帮助,河北省电力局用电处的东明辰先生和调通局的张英怀先生提供了预测的实际经验,他们提出了许多宝贵意见,华北电力大学邢棉同志提供了部分理论问题。在此我们表示深切的谢意。

华北电力大学陈志业教授认真审阅了全稿,提高了全书质量,在此表示衷心地感谢。

由于写作时间仓促,作者水平及经验所限,书中不足之处在所难免,恳请读者批评指正。

作 者

1998年4月

目 录

前 言

第一章 电力负荷预测概论	1
第一节 负荷预测概念和原理	1
第二节 用电量与负荷预测的分类	5
第三节 负荷预测基本程序	6
第四节 负荷预测误差分析	9
第二章 负荷预测经验技术与经典技术	16
第一节 专家预测法	16
第二节 类比法	17
第三节 主观概率预测法	17
第四节 单耗法	18
第五节 负荷密度法	18
第六节 比例系数增长法	19
第七节 弹性系数法	19
第三章 电力负荷趋势外推预测技术	25
第一节 水平趋势预测技术	25
第二节 线性趋势预测技术	31
第三节 多项式趋势预测技术	37
第四节 季节型趋势预测技术	44
第五节 增长趋势预测技术	54
第四章 电力负荷回归模型预测技术	61
第一节 一元线性回归模型及其参数估计	61
第二节 一元线性回归模型的检验	65
第三节 一元线性回归模型的预测分析	68
第四节 多元线性回归预测模型	73
第五节 非线性回归预测模型	86
第六节 线性回归模型在电力负荷预测中的应用	90
第五章 电力负荷时间序列预测技术	97
第一节 时间序列的概念	97
第二节 时间序列的线性模型	100
第三节 线性模型的自相关函数和偏相关函数	104
第四节 模型识别与参数估计	119
第五节 模型的检验	134
第六节 非平稳序列模型	137
第七节 随机模型在电力负荷预测中的应用	145
第六章 电力负荷灰色预测技术	165
第一节 灰色系统理论介绍	165

第二节 灰色生成	168
第三节 灰色建模过程	171
第四节 负荷灰色预测技术的改进	180
第七章 电力负荷预测技术的新发展	188
第一节 优选组合预测技术	188
第二节 专家系统预测技术	196
第三节 神经网络预测技术	199
第四节 小波分析预测技术	204
第八章 负荷预测系列软件介绍	210
第一节 用电负荷预测系统	210
第二节 日负荷预测系统	213
第三节 日周月负荷预测系统	214
第四节 重大节日负荷预测系统	215
附录一 矩阵及其运算	216
附录二 概率论基本知识	219
参考文献	231

第一章 电力负荷预测概论

本章介绍了电力负荷预测的基本概念和作用,论述了负荷预测常用方法的分类,介绍了负荷预测的误差分析手段,并给出了负荷预测工作的完整步骤,从而使读者对负荷预测理论方法达到初步了解,为后续章节的学习打好基础。

第一节 负荷预测概念和原理

电力系统的任务是给广大用户不间断地提供优质电能,满足各类负荷的需求。

一、负荷预测概念

负荷可指电力需求量或者用电量,而需求量是指能量的时间变化率,即功率。也可以说,负荷是指发电厂、供电地区或电网在某一瞬间所承担的工作负荷。对用户来说,用电负荷是指连接在电网的用户所有用电设备在某一瞬间所消耗的功率之和。

1. 负荷按物理性能划分

负荷按物理性能分为有功负荷和无功负荷。

(1) 有功负荷:是把电能转换为其他能量,并在用电设备中真实消耗掉的能量,计算单位为 kW (千瓦)。

(2) 无功负荷:在电能输送和转换过程中,需要建立磁场(如变压器、电动机等)而消耗的功率。它仅完成电磁能量的相互转换,并不做功,在这个意义上称为“无功”,计算单位是 kvar (千乏)。

2. 负荷按电能的划分

负荷按电能的产、供、销生产过程分为发电负荷、供电负荷和用电负荷。

(1) 发电负荷:指某时刻电网或发电厂的实际发电出力的总和,计算单位为 kW。

(2) 供电负荷:指供电地区内各发电厂发电负荷之和,减去发电及供热厂用电负荷,加上从供电地区外输入的负荷,再减去向供电地区外输出的负荷,计算单位为 kW。

(3) 用电负荷:指地区供电负荷减去线路和变压器中的损耗后的负荷,计算单位为 kW。

3. 负荷按时间的划分

负荷按时间分为年、月、日、时、分负荷。

4. 售电量及用电量

(1) 售电量:是指电力企业售给用户(包括趸售户)的电量及供给本企业非电力生产(如修配厂用电)、基本建设、大修理和非生产部门(如食堂、宿舍)等所使用的电量。

(2) 用电量:是指电网(或电力企业)的售电量与自备电厂自发、自用电和其售给附近用户的电量之和。

5. 电量的划分

电量可分为有功电量和无功电量。

(1) 有功电量:是指有功负荷与时间的乘积。有功电量可由电能表读出,也可由有功负

荷的平均值乘以时间得出,有功电量的计算单位是 $\text{kW} \cdot \text{h}$ 。

(2) 无功电量:是指无功负荷与时间的乘积。无功电量可由无功电能表读出,也可由无功负荷的平均值乘以时间得出,无功电量的计算单位是 $\text{kvar} \cdot \text{h}$ 。

6. 负荷预测

在充分考虑一些重要的系统运行特性、增容决策、自然条件与社会影响的条件下,研究或利用一套系统地处理过去与未来负荷的数学方法,在满足一定精度要求的意义下,确定未来某特定时刻的负荷数值,称为负荷预测。

二、电力负荷预测的作用和现状

电力负荷预测是供电部门的重要工作之一,准确的负荷预测,可以经济合理地安排电网内部发电机组的启停,保持电网运行的安全稳定性,减少不必要的旋转储备容量,合理安排机组检修计划,保证社会的正常生产和生活,有效地降低发电成本,提高经济效益和社会效益。

负荷预测的结果,还可以有利于决定未来新的发电机组的安装,决定装机容量的大小、地点和时间,决定电网的增容和改建,决定电网的建设和发展。

因此,电力负荷预测工作的水平已成为衡量一个电力企业的管理是否走向现代化的显著标志之一,尤其在我国的电力事业空前发展的今天,用电管理走向市场,电力负荷预测问题的解决已经成为我们面临的重要而又艰巨的任务。

负荷预测的核心问题是预测的技术方法,或者说是预测数学模型,随着现代科学技术的不断进步,负荷预测理论技术得到了很大发展,理论研究逐步深入,适合本地特点的预测程序、软件开始出现。如华北电力大学的陈志业、牛东晓教授等人先后对负荷预测理论及应用进行了研究,开发了适合短、中、长期各类负荷预测的实用软件包,分别通过了电力工业部和有关网省局的技术鉴定,鉴定认为负荷预测模型的研究达到了国际先进水平,并已广泛地应用于华北电网各个地区。

三、影响负荷预测作用大小的因素

负荷预测作用的大小,要看由于使用了预测结果所产生的收益,是否超出了所支出的费用,以及超出多少。影响预测作用大小的因素是多方面的,主要有以下几项。

1. 负荷预测费用的高低

负荷预测费用包括设计和实行预测程序费,历史数据资料收集、整理、计算和储存费,资料使用费,资料更新费,人员技术培训费等。

2. 负荷预测结果的准确性

一般说来,准确性高的负荷预测比准确性低的预测作用更大。但是,准确性高的预测方法往往是比较复杂的,这又与较大的费用支出相矛盾。这就要看由于使用复杂预测技术手段所提高的预测准确性,给决策者带来的好处有多大,是否值得。不同的负荷与电量预测对准确性的要求不同,长期的负荷预测甚至容许误差达到 10%,而短期的日负荷预测的误差一般不能超过 3%。

3. 负荷预测的时效性

所谓负荷预测的时效性指的是提出一项预测结果需要多少时间。如能很快地得到预测结果,可使决策者有充分的时间改变决策,即预测的领先时间长,则预测作用大。反之,如果迟迟拿不出预测结果,领先时间又很短,其作用也就不大了。这一点对于电力系统在线超短

期实时预测就更为重要了。为此,我们应该大力提倡使用现代科学技术手段,如高速大容量电子计算机,各类资料信息管理系统,获取气象、政治、经济信息的现代化通信手段。

4. 负荷预测所依据的历史资料其变动规律有无重大变化

在利用历史资料进行外推负荷预测中,如果负荷的过去和现在的发展规律直接延伸到未来,没有什么重大的干扰和破坏,则可以加以模型化,利用已知的模型,类比现在,预测未来。如果在预测期中发生了无法估计的重大事件(如气象的剧烈变化、严重灾害、国家政策的重大变化等),以致使负荷变化的正常规律被破坏,使原来持续上升或下降的资料发生转折,就会使预测失效。这类负荷预测失实的情况不少,因为转折点是最难预测的。

5. 负荷预测期限的长短

负荷预测的具体作用视预测期限的长短而异。电力负荷预测一般可分为长期、中期、短期和超短期四种,它们各有其预测的目的和意义。长期预测一般指10年以上,中期预测一般指5年左右,短期预测一般指1天以上至1年以下,超短期预测一般指1h以内。不同时期的预测会对不同的研究指标起不同的作用。

四、负荷预测的特点

由于负荷预测是根据电力负荷的过去和现在推测它的未来数值,所以,负荷预测工作所研究的对象是不肯定事件。只有不肯定事件、随机事件,才需要人们采用适当的预测技术,推知负荷的发展趋势和可能达到的状况。这就使负荷预测具有以下明显的特点。

1. 不准确性

因为电力负荷未来的发展是不肯定的,它要受到多种多样复杂因素的影响,而且各种影响因素也是发展变化的。人们对于这些发展变化有些能够预先估计,有些却很难事先预见到,加上一些临时情况发生变化的影响,因此就决定了预测结果的不准确性或不完全准确性。

2. 条件性

各种负荷预测都是在一定条件下作出的。对于条件而言,又可分为必然条件和假设条件两种,如果负荷员真正掌握了电力负荷的本质规律,那么预测条件就是必然条件,所作出的预测往往是比较可靠的。而在很多情况下,由于负荷未来发展的不肯定性,所以就是一些假设条件。例如,我们经常说,如果天气一直不下雨的话,排灌负荷将保持较高的数值等等。当然,这些假设条件不能毫无根据的凭空假设,而应根据研究分析,综合各种情况而得来。给预测结果加以一定的前提条件,更有利于用电部门使用预测结果。

3. 时间性

各种负荷预测都有一定的时间范围,因为负荷预测属于科学预测的范畴,因此,要求有比较确切的数量概念,往往需要确切地指明预测的时间。

4. 多方案性

由于预测的不准确性和条件性,所以有时要对负荷在各种情况下可能的发展状况进行预测,就会得到各种条件下不同的负荷预测方案。

五、负荷预测的基本原理

负荷预测工作是根据电力负荷的发展变化规律,预计或判断其未来发展趋势和状况的活动,因此必须科学地总结出预测工作的基本原理,用于指导负荷预测工作。

1. 可知性原理

也就是说,预测对象的发展规律,其未来的发展趋势和状况是可以为人们所知道的。客

观世界是可以被认识的，人们不但可以认识它的过去和现在，而且可以通过总结它的过去和现在推测其未来。这是人们进行预测活动的基本依据。

2. 可能性原理

因为事物的发展变化是在内因和外因共同作用下进行的。内因的变化及外因作用力大小不同，会使事物发展变化有多种可能性。所以，对某一具体指标的预测，往往是按照其发展变化的多种可能性，进行多方案预测的。

3. 连续性原理

又称惯性原理，是指预测对象的发展是一个连续统一的过程，其未来发展是这个过程的继续。它强调了预测对象总是从过去发展到现在，再从现在发展到未来。它认为事物发展变化过程中会将某些原有的特征保持下来，延续下去。电力系统的发展变化同样存在着惯性，如某些负荷指标会以原有的趋势和变化率发展下去。这种惯性正是我们进行负荷预测的主要依据。因此，了解事物的过去和现在，并掌握其变化规律，就可以对其未来的发展情况利用连续性原理进行预测。

4. 相似性原理

尽管客观世界中各种事物的发展各不相同，但一些事物发展之间还是存在着相似之处，人们就利用这种相似性进行预测。在很多情况下，作为预测对象的一个事物，其现在的发展过程和发展状况可能与另一事物过去一定阶段的发展过程和发展状况相类似，人们就根据后一事物的已知发展过程和状况，来预测所预测对象的未来发展过程和状况，这就是相似性原理。目前，预测技术中使用的类推法或历史类比法，就是基于这个原理的预测方法。例如，当我们预测一个新的经济开发区的用电量时，由于其建成时期较短，没有很多历史数据可利用，这时，就难以用趋势外推、回归分析等方法建模预测。这种情况下，我们可以参考一个早已建成的、规模和条件具有可比性的其他经济开发区，以其发展时期相对应的用电量，作为预测新经济开发区用电量的基础，从而可以作出相应的预测结果。

5. 反馈性原理

反馈就是利用输出返回到输入端，再调节输出结果。预测的反馈性原理实际上是为了不断提高预测的准确性而进行的反馈调节。人们在预测活动实践中发现，当预测的结果和经过一段实践所得到的实际值存在着差距时，可利用这个差距，对远期预测值进行反馈调节，以提高预测的准确性。在进行反馈调节时，首先认真分析预测值和实际值之间的差距及产生差距的原因，然后根据已经查明原因，适当改变输入数据，进行反馈，调节远期预测结果。反馈性预测实质上就是将预测的理论值与实际相结合，在实践中检验，然后进行修改、调整，使预测质量进一步提高。

6. 系统性原理

这个原理认为预测对象是一个完整的系统，它本身有内在的系统，它与外界事物的联系又形成了它的外在系统。这些系统综合成一个完整的总系统，都要进行考虑。即预测对象的未来发展是系统整体的动态发展，而且整个系统的动态发展与它的各个组成部分和影响因素之间的相互作用和相互影响密切相关。系统性原理还强调系统整体最佳，只有系统整体最佳的预测，才是高质量的预测，才能为决策者提供最佳的预测方案。

第二节 用电量与负荷预测的分类

用电分类是说明国民经济各部门用电情况和变化规律，它是反映电气化的发展水平和趋势的指标，用于分析研究经济增长与电力生产增长、社会产品增长与电力消耗量增长的相互关系，是负荷预测和电力分配的依据。为了统一历史数据的口径，在此做一介绍。

我国用电分类从 1986 年起按“国民经济行业用电分类”，把全部用电量按下列三个原则进行划分：一是参考国际行业分类的标准和经验，从我国实际情况出发划分各行业的界限，并在具体分类中兼顾国际资料对比的需要；二是区分物质生产领域和非物质生产领域；三是主要按照企业、事业单位、机关团体和个体从业人员所从事的生产或其他社会活动性质的同一性分类，即按其所属行业分类，而不按其所属行政主管部门系统分类，但也应适当照顾行政主管部门业务管理范围的需要。

一、国民经济行业用电分类

国民经济行业按用电分类，连同城乡居民用电可分为八大类，各类用电包括范围如下。

1. 农、林、牧、渔、水利业

该行业用电指乡、村用于农业、林业、畜牧业、渔业、水利业等方面的生产用电。

- (1) 农业包括种植业、其他农业；
- (2) 林业包括林业经营种植等；
- (3) 畜牧业包括牲畜饲养放牧业、家禽饲养业、其他畜牧业；
- (4) 渔业包括海水渔业和淡水渔业；
- (5) 水利业；
- (6) 农、林、牧、渔、水利服务业。

2. 工业

工业用电指工业企业用于工业生产的用电。如果企业生产跨行业产品时，应按企业主要产品进行归类。工业包括采掘业和制造业。

(1) 采掘业。采掘业主要包括矿业、木材及竹材采运业以及自来水生产和供应业。其中矿业包括固体矿藏、液体矿藏和气体矿藏的开采和洗选。主要有煤炭、石油和天然气、黑色金属、有色金属、建材及其他非金属矿的采选业、采盐业和其他矿采选业。

(2) 制造业。制造业范围较为广泛，包括 17 类：①食品、饮料和烟草制造业；②纺织业；③造纸业；④电力、蒸汽、热水生产和供应；⑤石油加工业；⑥煤焦、煤气及煤制品业；⑦化学工业中的轻工业；⑧医药工业；⑨化纤工业；⑩橡胶及塑料制品业中的轻工业；⑪建筑材料及其他非金属矿制品业；⑫黑色金属冶炼及压延加工业；⑬有色金属冶炼业；⑭金属制品业；⑮机械工业；⑯交通运输、电气、电子设备制造业；⑰其他工业。

3. 其他事业

- (1) 房地产管理、公用事业、居民服务和咨询服务业。
- (2) 卫生、体育和社会福利事业。
- (3) 教育、文化艺术和广播电视事业。
- (4) 科学研究和综合技术服务事业。
- (5) 国家党政机关和社会团体。

4. 城乡居民用电

(1) 乡村居民生活照明用电, 包括照明、家用电器等用电。

(2) 城镇居民生活照明用电, 包括照明、家用电器等用电。

此外, 国民经济用电还包括地质普查和勘探业, 建筑业, 交通运输邮电通信业, 商业、公共饮食业, 物质供销和仓储业。

二、负荷预测分类

1. 负荷预测按时间分类

电力负荷预测中经常按时间期限进行分类, 通常分为长期、中期、短期和超短期负荷预测。长期负荷预测一般指 10 年以上并以年为单位的预测, 中期指 5 年左右并以年为单位的预测。它们的意义在于帮助决定新的发电机组的安装 (包括装机容量大小、型式、地点和时间) 与电网的规划、增容和改建, 是电力规划部门的重要工作之一。

短期预测则是指一年之内以月为单位的负荷预测, 还指以周、天、小时为单位的负荷预测, 通常预测未来一个月度、未来一周、未来一天的负荷指标, 也预测未来一天 24h 中的负荷。其意义在于帮助确定燃料供应计划; 对运行中的电厂出力要求提出预告, 使对发电机组出力变化事先得以估计; 可以经济合理地安排本网内各机组的启停, 降低旋转储备容量; 可以在保证正常用电的情况下合理安排机组检修计划。

超短期负荷预测指未来 1h、未来 0.5h 甚至未来 10min 的预测。其意义在于可对电网进行计算机在线控制, 实现发电容量的合理调度, 满足给定的运行要求, 同时使发电成本最小。

2. 负荷预测按行业分类

负荷预测可以分为城市民用负荷、商业负荷、农村负荷、工业负荷以及其他负荷的负荷预测。其中, 城市民用负荷预测主要指城市居民的家用负荷预测; 商业负荷预测和工业负荷预测是指对各自为商业与工业服务的负荷进行预测; 农村负荷预测是指广大农村所有负荷 (包括农村民用电、生产与排灌用电以及商业用电等) 的预测; 而其他负荷预测则包括市政用电 (如街道照明等)、公用事业、政府办公、铁路与电车、军用等等负荷的预测。

虽然负荷可以大致这样分类预测, 但并不严格, 对于按某类负荷进行预测时, 可能发生的某些实际负荷归算到哪一类负荷的争执。在这种情况下, 就只能由各供电部门自己决定。因此, 在一些供电公司中, 可以各自有其更具体的负荷预测分类细目。

3. 负荷预测按特性分类

根据负荷预测表示的不同特性, 常常又分为最高负荷、最低负荷、平均负荷、负荷峰谷差、高峰负荷平均、低谷负荷平均、平峰负荷平均、全网负荷、母线负荷、负荷率等类型的负荷预测, 以满足供电、用电部门的管理工作的需要。

第三节 负荷预测基本程序

对电力负荷进行科学预测, 要有一个基本程序, 就是要考虑预测工作怎样进行, 分几个阶段, 先做什么, 后做什么。只有把负荷预测工作的整个程序搞清楚, 才能做好负荷预测工作。根据所进行的电力负荷预测的实践活动, 认为其基本程序如下。

一、确定负荷预测目的, 制订预测计划

负荷预测目的要明确具体, 紧密联系电力工业实际需要, 并据以拟定一个负荷预测工作

计划。在预测计划中要考虑的问题主要有：准备预测的时期，所需要的历史资料（按年、按季、按月、按周或按日），需要多少项资料，资料的来源和搜集资料的方法，预测的方法，预测工作完成时间，所需经费来源等等。关于所需资料项数多少，说法不一。有人主张外推预测的时期数不能超过历史资料的时期数，如设 d = 历史资料时期数， h = 外推预测时期数，则有 $d \geq h$ 。也有人认为，这种要求低估了短期预测所需项数和高估了长期预测所需项数，主张用 $d = 4\sqrt{h}$ 计算。按此式，如向前预测 1 期，则 $h = 1$ ， $d = 4$ ，即需要 4 期历史资料；如向前预测 4 期，则需 8 期历史资料；如向前预测 100 期，就要用 40 期历史资料即可。可见，用这个公式计算，照顾了短期预测的需要，却不利于长期预测。实际上，根据长期的历史资料进行短期预测，要比根据短期的历史资料进行长期预测更可靠些，因为这样根据更充分些。

二、调查资料 and 选择资料

要多方面调查收集资料，包括电力企业内部资料和外部资料，国民经济有关部门的资料，以及公开发表和未公开发表的资料，然后从众多的资料中挑选出有用的一小部分，即把资料浓缩到最小量。挑选资料的标准，一要直接有关性，二要可靠性，三要最新性。先把符合这三点的资料挑出来，加以深入研究，在这以后，才能考虑是否还需要再收集其他资料。收集统计资料是不容易的，尤其是在我国当前的情况下，各层次的资料往往不够完整，真实性也有问题，再加上保密问题尚未解决，就更增加了难度。尤其是如果资料收集和选择得不好，会直接影响负荷预测的质量。

三、资料整理

对所收集的与负荷有关的统计资料进行审核和必要的加工整理，是保证预测质量所必须的。可以说，预测的质量不会超过所用资料的质量，整理资料的目的是为了保证资料的质量，从而为保证预测质量打下基础。

1. 衡量统计资料质量的标准

衡量一个统计资料质量高低的标准，主要有以下几方面：

- (1) 资料完整无缺，各期指标齐全；
- (2) 数字准确无误，反映的都是正常（而不是反常）状态下的水平，资料中没有异常的“分离项”（outlier）；
- (3) 时间数列各值间有可比性。

此外，还有历史资料的表现形式是否适合需要，是否需要变换，以及计量单位是否规范化等问题也要注意。

2. 资料的整理

资料整理的主要内容有以下几项：

- (1) 资料的补缺推算。如果中间某一项的资料空缺，则可利用相邻两边资料取平均值近似代替，如果开头一项资料空缺，则可利用趋势比例计算代替。
- (2) 对不可靠的资料加以核实调整。对能查明原因的异常值，用适当方法加以订正；对原因不明而又没有可靠修改根据的资料，最好删去，有多少删多少。
- (3) 对时间数列中不可比资料加以调整。时间数列资料的可比性主要包括：各期统计指标的口径范围是否完全一致；各期价值指标所用价格有无变动；各期时间单位长度是否可比；周期性的季节变动资料的各期资料是否可比，是否能如实反映周期性变动规律。

用不同方法处理上述各种可比性问题时，务必使资料在时间上有可比性。此外，还要根

据研究目的,认真考虑时间数列的起止时间,即应截取哪一段时期的资料使用。

四、对资料的初步分析

在经过整理之后,还要对所用资料进行初步分析,包括以下几方面:

(1) 画出动态折线图或散点图,从图形中观察资料变动的轨迹,特别注意离群的数值(异常值)和转折点,研究它是由偶然的,还是其他什么确定的原因所致。

(2) 查明异常值的原因后,加以处理,对于异常值,常用的处理方法是,设负荷历史数据为 $x_1, \dots, x_n$, 令 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, 若 $x_i > \bar{x} (1+20\%)$, 取 $x_i = \bar{x} (1+20\%)$; 若 $x_i < \bar{x} (1-20\%)$, 取 $x_i = \bar{x} (1-20\%)$ 。从而使历史数据序列趋于平稳。

除此之外,也有用非平稳序列的平稳化代换方法和灰色系统的累加生成方法进行处理的,这些技术将在后面讲到。

(3) 计算一些统计量,如自相关系数,以进一步辨明资料轨迹的性质,为建立模型做准备。这些统计量将在后面介绍。

五、建立预测模型

负荷预测模型是统计资料轨迹的概括,它反映的是经验资料内部结构的一般特征,与该资料的具体结构并不完全吻合。模型的具体化就是负荷预测公式,公式可以产出与观察值有相似结构的数值,这就是预测值。负荷预测模型是多种多样的,以适用于不同结构的资料,因此,对一个具体资料,就有选择适当预测模型的问题。正确选择预测模型在负荷预测中是具有关键性的一步。有时由于模型选择不当,造成预测误差过大,就需要改换模型。必要时,可同时采用几种数学模型进行运算,以便对比、选择。

六、综合分析,确定预测结果

通过选择适当的预测技术,建立负荷预测数学模型,进行预测运算得到的预测值,或利用其他方法得到的初步预测值,还要参照当前已经出现的各种可能性,以及新的趋势与发展,进行综合分析、对比、判断推理和评价,最终对初步预测结果进行调整和修正。这是因为从过去到现在的发展变化规律,不能说就是将来的变化规律。所以要对影响预测对象的新因素进行分析,对预测模型进行适当的修正后确定预测值。

预测值的确定决不是通过某一、两个预测运算就能轻而易举取得的。搞好预测需“重在分析、贵在方法、巧在应用”,也就是说,负荷预测工作不仅是一种科学,而且是一种艺术,良好的综合判断能力是难于用简单的语言传授的,而是个人才能、经验与教训综合作用的结果。

七、编写预测报告,交付使用

根据分析判断最后确定的预测结果,编写出本次负荷预测的报告。因为预测结果经常是多方案的,所以报告中要对取得这些结果的预测条件、假设及限制因素等情况详细说明。在报告中应有数据资料、报告分析、数学模型、预测结果及必要的图表,让使用者一目了然,便于应用。

八、负荷预测管理

将负荷预测报告提交主管部门后,只是本次预测报告一段落,并不等于全部预测工作的结束,随后仍需根据主客观条件的变化及预测应用的反馈信息进行检验,必要时应修正预测值。例如,预测值交付使用后,经过一段时间的实践,发现这一时期的实际值和预测值之间有差

距,就要利用反馈性原理对远期预测值进行调整,这也是对负荷预测的滚动性管理。

对预测结果还要进行预测误差分析,如果从分析中发现预测误差偏大,就要检查原因,是不是影响历史负荷变动的基本因素发生了变化,以致负荷的轨迹变了,从而考虑改换模型。对误差数列的分析有助于辨明所拟合的模型是否充分,是否适当。

第四节 负荷预测误差分析

由于负荷预测是一种对未来负荷的估算,因此,它与客观实际还是存在着一定的差距,这个差距就是预测误差。

预测误差和预测结果的准确性关系密切。误差愈大,准确性就愈低;反之,误差愈小,准确性就愈高。可见,研究产生误差的原因,计算并分析误差的大小,是有很意义的。这不但可以认识预测结果的准确程度,从而在利用预测资料作决策时具有重要的参考价值,同时,对于改进负荷预测工作,检验和选用恰当的预测方法等方面也有很大帮助。

一、产生误差原因

产生预测误差的原因很多,主要有以下几个方面:

(1) 进行预测往往要用到数学模型,而数学模型大多只包括所研究现象的某些主要因素,很多次要的因素都被略去了。对于错综复杂的电力负荷变化来说,这样的模型只是一种经过简单化了的负荷状况的反映,与实际负荷之间存在差距,用它来进行预测,也就无可避免地会与实际负荷产生误差。

(2) 负荷所受影响是千变万化的,进行预测的目的和要求又各种各样,因而就有一个如何从许多预测方法中正确选用一个合适的预测方法的问题。如果选择不当的话,也就随之而产生误差。

(3) 进行负荷预测要用到大量资料,而各项资料并不能保证都是准确可靠,这就必然会带来预测误差。

(4) 某种意外事件的发生或情况的突然变化,也会造成预测误差。此外,由于计算或判断上的错误,如平滑常数的选择不妥,也会产生不同程度的误差。

以上各种不同原因引起的误差是混合在一起表现出来的,因此,当发现误差很大,预测结果严重失实时,必须针对以上各种原因逐一进行审查,寻找根源,加以改进。

二、预测误差分析

计算和分析预测误差的方法和指标很多,现主要介绍如下几种。

1. 绝对误差与相对误差

设 Y 表示实际值, $\hat{Y}$ 表示预测值,则称 $Y - \hat{Y}$ 为绝对误差,称 $\frac{Y - \hat{Y}}{Y}$ 为相对误差。有时相对误差也用百分数 $\frac{Y - \hat{Y}}{Y} \times 100\%$ 表示。这是一种直观的误差表示方法。在电力系统中作为一种考核指标而经常使用。

2. 平均绝对误差

$$MAE = \frac{1}{n} \sum_{i=1}^n |E_i| = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|$$

式中 MAE ——平均绝对误差；

E_i ——第 i 个预测值与实际值的绝对误差；

Y_i ——第 i 个实际负荷值；

$\hat{Y}_i$ ——第 i 个预测负荷值。

由于预测误差有正有负，为了避免正负相抵消，故取误差的绝对值进行综合并计算其平均数，这是误差分析的综合指标法之一。

3. 均方误差

$$MSE = \frac{1}{n} \sum_{i=1}^n E_i^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

式中 MSE ——均方差，其他符号同前。

均方误差是预测误差平方之和的平均数，它避免了正负误差不能相加的问题。是误差分析的综合指标法之一。

4. 均方根误差

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n E_i^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}$$

式中 $RMSE$ ——均方根误差，其他符号同前。

这是均方误差的平方根。由于对误差 E 进行了平方，加强了数值大的误差在指标中的作用，从而提高了这个指标的灵敏性，是一大优点，这也是误差分析的综合指标之一。

5. 标准误差

$$S_Y = \sqrt{\frac{\sum (Y_i - \hat{Y}_i)^2}{n - m}} \quad (i = 1, 2, \dots, n)$$

式中 S_Y ——预测标准误差；

n ——历史负荷数据个数；

m ——自由度，也就是变量的个数，即自变量和因变量的个数的总和。

6. 关联度误差分析

关联度是灰色系统理论提出的一种技术方法，是分析系统中各因素关联程度的方法，或者说关联程度量化的方法。

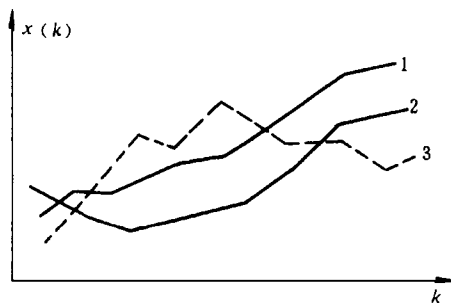


图 1-1 各因素几何态势图

关联度大，曲线 1、3 的关联度较小。

关联度的基本思想，是根据曲线间相似程度来判断关联程度，实质上是几种曲线间几何形状的分析比较，即认为几何形状越接近，则发展变化态势越接近，关联程度越大。用此方法可以来比较几种预测模型对应的几条预测曲线与一条实际曲线的拟合程度，关联度越大，则说明对应的预测模型越优，拟合误差也就越小。图 1-1 所示的各因素几何态势图，曲线 1、2 间的相似程度大于曲线 1、3 间的相似程度，因此认为曲线 1、2 的

如果说我们指定参考数列为 x_0 , 被比较数列 (又称预测数列或因素数列) 为 x_i , 其中 $i=1, 2, \dots, m$, 且

$$\begin{aligned}x_0 &= \{x_0(1), x_0(2), \dots, x_0(n)\} \\x_i &= \{x_i(1), x_i(2), \dots, x_i(n)\} \\i &= 1, 2, \dots, m\end{aligned}$$

则称

$$\xi_i(k) = \frac{\min_i \min_k |x_0(k) - x_i(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|}$$

为曲线 x_0 与 x_i 在第 k 点的关联系数。上式中: $|x_0(k) - x_i(k)| = \Delta_i(k)$ 称为第 k 点 x_0 与 x_i 的绝对差; $\min_i \min_k |x_0(k) - x_i(k)|$ 称为两级最小差, 其中 $\min_k |x_0(k) - x_i(k)|$ 是第一级最小差, 这表示在第 x_i 曲线上, 找各点与 x_0 的最小差, $\min_i \min_k |x_0(k) - x_i(k)|$ 是第二级最小差, 表示在各条曲线中找出的最小差基础上, 再按 $i=1, i=2, \dots, i=m$ 找所有曲线 x_i 中的最小差; $\max_i \max_k |x_0(k) - x_i(k)|$ 是两级最大差, 其意义与最小差相似; ρ 称为分辨系数, 是 0 与 1 间的数, 一般取 $\rho=0.5$ 。

综合各点的关联系数, 可得出整个 x_i 曲线与参考曲线 x_0 的关联度 r_i 为

$$r_i = \frac{1}{n} \sum_{k=1}^n \xi_i(k)$$

对于单位不同, 或初值不同的数列作关联度分析时, 首先要做无量纲化、归一化预处理, 也称为初值化。为了把所有数列无量纲化, 并且要求所有数列有公共交点, 用每一个数列的第一个数 $x_i(1)$ 除其他数 $x_i(k)$, 就可解决这两个问题, 使得各数列之间具有可比性。

【例 1-1】 设 y_0 为参考数列 (历史负荷数列), $\hat{y}_1$ 、 $\hat{y}_2$ 为比较数列 (两种预测方法所得预测值数列), y_0 、 $\hat{y}_1$ 、 $\hat{y}_2$ 分别为

$$\begin{aligned}y_0 &= \{y_0(1), y_0(2), y_0(3), y_0(4), y_0(5), y_0(6)\} \\&= \{20, 22, 40, 45, 60, 80\} \\ \hat{y}_1 &= \{\hat{y}_1(1), \hat{y}_1(2), \hat{y}_1(3), \hat{y}_1(4), \hat{y}_1(5), \hat{y}_1(6)\} \\&= \{30, 35, 55, 60, 70, 90\} \\ \hat{y}_2 &= \{\hat{y}_2(1), \hat{y}_2(2), \hat{y}_2(3), \hat{y}_2(4), \hat{y}_2(5), \hat{y}_2(6)\} \\&= \{40, 45, 43, 55, 65, 70\}\end{aligned}$$

试求 $\hat{y}_1$ 、 $\hat{y}_2$ 分别对 y_0 的关联度。

解: 第一步, “初值化”

$$\begin{aligned}x_0 &= \left\{ \frac{20}{20}, \frac{22}{20}, \frac{40}{20}, \frac{45}{20}, \frac{60}{20}, \frac{80}{20} \right\} \\&= \{1, 1.1, 2, 2.25, 3, 4\} \\x_1 &= \left\{ \frac{30}{30}, \frac{35}{30}, \frac{55}{30}, \frac{60}{30}, \frac{70}{30}, \frac{90}{30} \right\} \\&= \{1, 1.166, 1.833, 2, 2.333, 3\} \\x_2 &= \left\{ \frac{40}{40}, \frac{45}{40}, \frac{43}{40}, \frac{55}{40}, \frac{65}{40}, \frac{70}{40} \right\}\end{aligned}$$

$$=\{1, 1.125, 1.075, 1.375, 1.625, 1.75\}$$

第二步, 求关联系数中两级差

记 $\Delta_1(k)$ 为

$$\begin{aligned}\Delta_1(k) &= |x_0(k) - x_1(k)| \\ \Delta_1 &= \{\Delta_1(1), \Delta_1(2), \Delta_1(3), \Delta_1(4), \Delta_1(5), \Delta_1(6)\} \\ \Delta_1(1) &= |x_0(1) - x_1(1)| \\ &= |1 - 1| = 0 \\ \Delta_1(2) &= |x_0(2) - x_1(2)| \\ &= |1.1 - 1.166| = 0.066 \\ \Delta_1(3) &= |x_0(3) - x_1(3)| \\ &= |2 - 1.334| = 0.666 \\ \Delta_1(4) &= 0.25 \\ \Delta_1(5) &= 0.686 \\ \Delta_1(6) &= 1\end{aligned}$$

又

$$\begin{aligned}\Delta_2(k) &= |x_0(k) - x_2(k)| \\ \Delta_2 &= \{\Delta_2(1), \Delta_2(2), \Delta_2(3), \Delta_2(4), \Delta_2(5), \Delta_2(6)\} \\ &= \{0, 0.025, 0.925, 0.875, 1.375, 2.25\}.\end{aligned}$$

则

$$\begin{aligned}\min_k |x_0(k) - x_1(k)| &= \min_k \{\Delta_1(k)\} \\ &= \min\{\Delta_1(1), \Delta_1(2), \Delta_1(3), \Delta_1(4), \Delta_1(5)\} \\ &= \min\{0, 0.066, 0.666, 0.25, 0.686, 1\} \\ &= 0 \\ \min_i \min_k |x_0(k) - x_i(k)| &= 0 \\ \max_k |x_0(k) - x_1(k)| &= \max_k \{\Delta_1(k)\} \\ &= \max\{0, 0.066, 0.666, 0.25, 0.686, 1\} \\ &= 1 \\ \max_k |x_0(k) - x_2(k)| &= \max_k \{\Delta_2(k)\} \\ &= \max\{0, 0.025, 0.925, 0.875, 1.375, 2.25\} \\ &= 2.25 \\ \max_i \max_k |x_0(k) - x_i(k)| &= \max\left\{\max_k |x_0(k) - x_1(k)|, \max_k |x_0(k) - x_2(k)|\right\} \\ &= \max\{1, 2.25\} = 2.25\end{aligned}$$

第三步, 求关联系数

$$\xi_1(k) = \frac{\min_i \min_k |x_0(k) - x_i(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_1(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|}$$

由于, $|x_0(k) - x_1(k)| = \Delta_1(k)$, 并取 $\rho = 0.5$, 将两级差代入上式, 可求得两级最小差

$$\min_i \min_k |x_0(k) - x_i(k)| = 0$$

$$\max_i \max_k |x_0(k) - x_i(k)| = 2.25$$

由此可得

$$\xi_1(k) = \frac{0.5 \times 2.25}{\Delta_1(k) + 0.5 \times 2.25} = \frac{1.125}{\Delta_1(k) + 1.125}$$

$$\xi_1(1) = \frac{1.125}{\Delta_1(1) + 1.125} = \frac{1.125}{0 + 1.125} = 1$$

$$\begin{aligned} \xi_1(2) &= \frac{1.125}{\Delta_1(2) + 1.125} = \frac{1.125}{0.066 + 1.125} \\ &= 0.9445 \end{aligned}$$

$$\begin{aligned} \xi_1(3) &= \frac{1.125}{\Delta_1(3) + 1.125} = \frac{1.125}{0.166 + 1.125} \\ &= 0.8714 \end{aligned}$$

$$\begin{aligned} \xi_1(4) &= \frac{1.125}{\Delta_1(4) + 1.125} = \frac{1.125}{0.25 + 1.125} \\ &= 0.818 \end{aligned}$$

$$\begin{aligned} \xi_1(5) &= \frac{1.125}{\Delta_1(5) + 1.125} = \frac{1.125}{0.66 + 1.125} \\ &= 0.63 \end{aligned}$$

$$\begin{aligned} \xi_1(6) &= \frac{1.125}{\Delta_1(6) + 1.125} = \frac{1.125}{1 + 1.125} \\ &= 0.5294 \end{aligned}$$

即

$$\begin{aligned} \xi_1 &= \{\xi_1(1), \xi_1(2), \xi_1(3), \xi_1(4), \xi_1(5), \xi_1(6)\} \\ &= \{1, 0.9445, 0.8714, 0.818, 0.63, 0.5294\} \end{aligned}$$

相似地有

$$\xi_2(k) = \frac{1.125}{\Delta_2(k) + 1.125}$$

$$\xi_2(1) = \frac{1.125}{0 + 1.125} = 1$$

$$\xi_2(2) = \frac{1.125}{0.025 + 1.125} = 0.978$$

$$\xi_2(3) = \frac{1.125}{0.925 + 1.125} = 0.5487$$

$$\xi_2(4) = \frac{1.125}{0.875 + 1.125} = 0.5625$$

$$\xi_2(5) = \frac{1.125}{1.375 + 1.125} = 0.45$$

$$\xi_2(6) = \frac{1.125}{2.25 + 1.125} = 0.333$$

即

$$\begin{aligned} \xi_2 &= \{\xi_2(1), \xi_2(2), \xi_2(3), \xi_2(4), \xi_2(5), \xi_2(6)\} \\ &= \{1, 0.978, 0.5487, 0.5625, 0.45, 0.333\} \end{aligned}$$

第四步, 求关联度

$$r_1 = \frac{1}{n} \sum_{k=1}^n \xi_1(k) = \frac{1}{6} \sum_{k=1}^6 \xi_1(k)$$