

计算机在电网调度运行中 应用的新进展

(译文及国外情况汇编)

韩祯祥 孙左一 主 编

电力系统自动化杂志社

计算机在电网调度运行中应用的新进展

(译文及国外情况汇编)

主 编 韩祯祥 孙左一
译 校 者 邱家驹 于尔铿 查仁柏 顾锦汶
童建中 韩海航 郑纪蛟 魏光耀
夏祖治 于 渤 赵祖康 陈 堂
责任编辑 綦兵兵

电 力 系 统 自 动 化 杂 志 社

1989年1月

编 辑 说 明

1988年5月,浙江大学韩祯祥教授来信,建议在《电力系统自动化》杂志上连载《Proceedings of the IEEE》(Vol.75, No.12, December 1987)上刊出的“Computers in Power System Operation”专辑中的九篇文章。由Felix F. Wu和Ralph D. Masiello主持编辑的这一专辑全面反映了近年间国外电网调度运行中计算机应用的新进展,对国内电力行业有一定的参考价值。由于《电力系统自动化》杂志篇幅有限,安排这一连载有一定困难。同年8月,能源部在山东烟台召开全国电网调度自动化工作会议。会议期间,我向能源部调度通信局蔡洋副局长转达了韩祯祥教授的提议,蔡洋副局长在会议的间隙中浏览了专辑英文版后表示赞同韩祯祥教授的意见,希望我们能将这一组文章翻译出版,介绍给国内从事有关工作的同志们。正好这次会议委托电力系统自动化杂志编辑出版名为《中国电网调度自动化的现状与展望》的文献汇编,我们打算同时将这一组译文推出,使读者能够了解国内外的全面情况。

此议既定,韩祯祥教授亲自组织了浙江大学、浙江省电力局、电力科学研究院和中国电机工程学会的部分同志翻译这些稿子。调度通信局裴钟棣副总工程师又向我们推荐了英国CEGB A.B. Baker撰写的《负荷预测在大型电力系统发电计划制定中的应用》一文。另外, T.E. Dy Liacco 向国内有关部门赠送了他收集的包括我国在内的世界各国300个电网调度中心的计算机配置情况,亦一并提供收入,形成一套较全面完整的资料。他山之石,可以攻玉,相信读者会从中得到一些有益的启示。

参加全书译校工作的有邱家驹、于尔铿、查仁柏、顾锦汶、童建中、韩海航、郑纪蛟、魏光耀以及夏祖治、于渤、赵祖康、陈堂等。编辑工作并得到王积荣、陈开庸、李一中等同志的支持和帮助。

本书责任编辑是蔡兵兵。参加编辑出版工作的还有赵楚玮、郭敏、叶薇君等。

限于编辑者的水平,书中一定会有疏漏和差错之处,祈望读者予以指正。

出于印刷技术的考虑,本书参考文献及“世界各国电网调度中心计算机配置情况总览”作为附本另印成册。

电力系统自动化杂志社 孙左一

1988年12月于南京

目 录

短期负荷预测.....	(1)
运行计划的最优化方法.....	(18)
发电计划及控制.....	(39)
电力网络的实时建模.....	(57)
安全分析与最优化.....	(79)
监控和数据采集.....	(106)
计算机的配置.....	(124)
系统间的通信及联网.....	(138)
电力系统运行中的人工智能技术.....	(149)
负荷预测在大型电力系统发电计划制定中的应用.....	(159)
参考文献.....	(见附本)
世界各国电网调度中心计算机配置情况总览.....	(见附本)

短期负荷预测

SHORT-TERM LOAD FORECASTING

G. Gross F. D. Galiana

提 要

本文讨论从一小时至一周内的短期负荷预测(STLF)的最新发展动态。文章回顾了STLF在能源管理系统(EMS)在线发电计划和安全功能中的重要性,讨论了负荷的性质和不同因素对其行为的影响。提出了负荷模型类型的详细分类和预测技术。在介绍分类的同时,推荐了一些支持展开上述讨论的有关文献。文章还广泛地讨论了STLF模型和软件包的开发和应用的实际情况。附注的文献目录列举了STLF方面主要出版物中有代表性的著作。

一、引 言

系统发电出力随时紧跟系统负荷的变化是电力系统运行的基本要求。为了有效地经济运行和控制,这一点必须在很大的时段内都得到满足。在以秒计的时段内,负荷的变化是很小的,且是随机的。自动发电控制(AGC)的功能保证了实际发电出力等于负荷功率。在以分计的时段内,当发生可能的大负荷波动时,经济调度功能用来使负荷在投入运行的发电机组间进行经济分配。在以小时或天计的时间段内,负荷和变化幅度更大。为满足这时间段内负荷变化的需要,有时候必须启动和停止某些发电机组或改变与相邻系统的交换功率,这是由一系列发电控制的功能来决定的,如水利计划、开停机计划、水火电协调和交换功率的估算等。在以周计的时段内,当大幅度的负荷摆动发生时,如燃料、水能和检修计划等功能都被用来保证以现有的资源组合来最经济地满足负荷的需要。另外,为了保证电力系统未来的安全运行,需要用离线的网络分析功能来研究电力系统在各种假想事故条件下的行为。所有这些功能无例外地都需要知道系统的负荷数值。在实时的环境下,状态估计被用来检验远方量测的有效性,从而估计出每条母线电压的幅值和角度。这些数值可以用来计算即时的负荷估计数。超短期的负荷预计程序分别嵌套在AGC和经济调度功能之中,分别超前几秒和几分。用于水利计划、开停机安排、水火协调和交换功率估算功能的负荷信息是从短期负荷预测系统获得。燃料和用水分配及检修计划功能要求预计一周以上的负荷,这些负荷预计是由运行计划预测系统得到的。

(一) 定义和范围

本文是关于电力系统运行的短期负荷预测(STLF)方面的,全文中“短期”一词意思是以小时计的预测时间。时间范围从下一小时,如可能的话下半小时,直至168小时。STLF中感兴趣的基本的典型数值是各小时的系统总负荷。除系统各小时的负荷值外,STLF同时也预测如下各量:

1. 系统日负荷峰值;
2. 一天中某一时刻的系统负荷值;
3. 系统每小时或半小时的电量;
4. 系统一天或一周的电量。

在本文中,在STLF的范围内,不但包括每小时或半小时直至168小时的负荷预测值,同时也包括任一或所有这些数值(对那些系统负荷是以半小时计的系统,预测是按每半小时做的)。

(二) STLF的重要性

STLF在电力系统经济的、可靠的和安全的运行策略的形成中起关键的作用,STLF功能的原则目标是为下列方面提供负荷预测:

1. 基本的发电计划功能;
2. 随时的电力系统安全性分析;
3. 实时的调度员信息。

STLF功能的主要用途是驱使发电计划功能去决定最经济的电源开停计划,以求与可靠性要求、运行约束和政策以及物理的、环境的和设备的限制相一致。对纯水电系统而言,需要负荷预计为水电规划功能确定最优的水库放水和发电厂的出力水平;对纯火电系统而言,开停机计划功能需要负荷预测,以决定为供给预计负荷的最小费用的开停机安排;而对于水火混合系统,水火电协调功能需要负荷预测以计划每小时的各种电源的运行水平,以求最低的生产成本。上述水电计划或开停机计划或水火电协调功能都要求预计下一天或下一周的系统负荷以决定满足系统运行的各种约束条件的最小费用的运行计划。另一个息息相关的计划任务是根据交换功率估算功能规划和签订交换功率合同。为此,短期负荷预计同样被用来确定与相邻电力公司的经济的交换功率水平。

STLF的第二个用途是用于电力系统安全的预防性分析。系统负荷的预测是离线网络分析功能的基本数据要求,该功能用来推断未来运行的系统是否是脆弱的。这一信息容许调度员去准备必要的校正手段(例如开调峰机组、切除负荷、功率求援、倒闸操作)以使系统安全运行。

STLF的第三个用途是提供给系统调度员以及时的信息。例如最近的负荷预测量、最新的气象预报和考虑到的随机行为。调度员需要这些信息使系统经济地、可靠地运行。图1中归纳了STLF的主要用途。

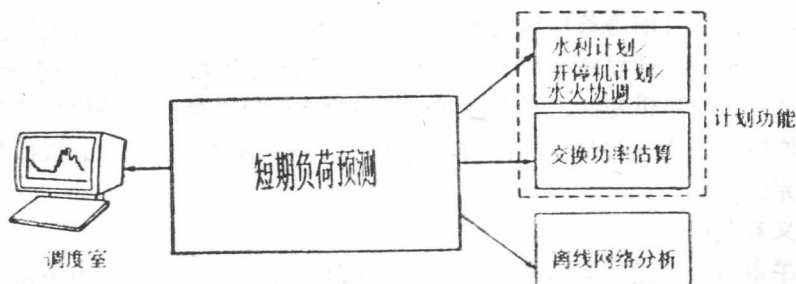


图1 短期负荷预测功能的主要用途是给调度员提供信息并把这
些信息输入到计划功能和离线安全分析中去

(三) EMS中的STLF

以前由电力系统调度员所作的人工预测已由现代能源管理系统中的STLF软件包所替代。STLF系统中的主要成份是STLF模型、数据来源和人——机对话(MMI)。STLF模型实现系统负荷的表示和STLF算法。数据来源有历史负荷、气象数据库、参数数据库、调度员人工输入以及从EMS的AGC功能和气象预报来的实时数据。图2表示了STLF功能的数据输入方式。人工输入的数据包括气象情况修改, 负荷预测参数数据或执行命令。STLF一般使用综合负荷的电量值(MWh)。实时数据库中的远方量测被AGC所使用, 以确定被“量测”的负荷, 这些负荷先按典型的方式被总加(使变得平滑)然后供STLF模型使用, 并向调度室和其它需要负荷测量的EMS功能(见图1), 提供STLF的输出。

短期负荷预测的及时性和准确度明显地影响着电力系统运行和电能成本。系统调度员必须预计系统负荷曲线的模式, 才能有足够的发电容量去满足负荷的需要。与此同时, 足够水平的运转备用和冷备用对缓解负荷预测和发电机组的可用性方面固有的不确定因素的产生的冲击是必需的。备用的代价是高昂的, 因为备用机组不能带足负荷而运行在它们最经济出力之下。运转备用和冷备用容量由电力系统运行所希望的安全和可靠性尺度来决定。因此, 减少预测误差 备用容量就可以减少而不致影响到系统可靠性和安全性, 这样运行费用也就降低了。

另外, 负荷预测误差会导致运行费用的增加。对负荷的估计不足就不能提供必要的备用容量, 到时候反而动用昂贵的调峰机因而使费用增加。反过来, 对负荷作过高的估计将启动过多的机组, 致使备用容量和运行费用不必要地增加。据1985年以大机组为主的英国电力系统估计, 负荷预测误差每增加1%, 则运行费用将相应增加1千万英镑。

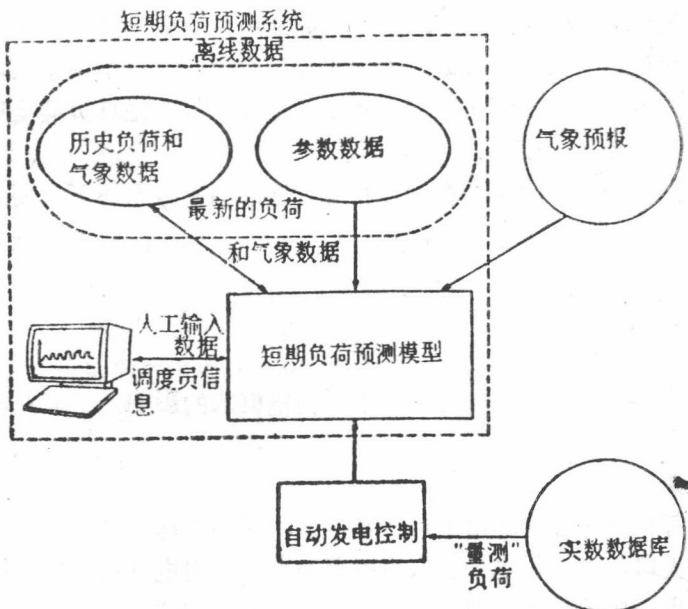


图2 短期负荷预测模型的输入数据来源

(四) 预测模型和技术

技术文献中罗列了各种STLF的方法和模型。因为没有两个电力公司是完全相同的, STLF模型从一个公司到另一个公司的可移植性也很有限。而另一方面, 出现在文献中的各

种技术——标准算法被裁剪成适用某系统特点或新开发的STLF程序却具有从一个公司到另一个公司的广泛的可移植性，本文综述有代表性的STLF模型和技术。

有几种密切相关的预测功能，如区域负荷和母线负荷预测。对那些跨越气候各异的地理区域或有气候差异的结构子单元（通称区域）的电力系统，区域负荷预测功能提供总的区域负荷预测量。区域短期负荷预测对于区域间联络线潮流的调整、区域发电计划的制订和母线负荷预测功能都是需要的。母线负荷预测是通过系统的或区域负荷预测功能给出关键母线的负荷预测值。母线负荷预测用于在线和离线的安全分析。本文不讨论区域和母线负荷预测，因为本文的焦点仅在短期系统负荷之上。但必须注意，这里讨论的许多STLF方法既可用于系统负荷也可用于母线和区域负荷。

（五）本文的要点

本文的目的是综观STLF并对其发展现状作出有代表性的评述。本文分五节。下一节我们将讨论系统负荷的性质，重点放在一个STLF模型中应予以考虑的原则性效应。接着讨论文献中各种STLF模型和预测程序，这些文献是根据模型的性质，数据和计算量以及预测要求来分类的。再下一节讨论STLF在控制中心环境下实现和应用的实际考虑。在结论一节中，勾划出STLF方面的某些未来可能的研究方向。全文引用的参考文献形成了以一系列STLF关键特性注释的文献目录的一部分。这不是一个全面的文献目录，我们向那些其作品未能包括在内或被误释了的作者表示歉意。不管怎么说，文献目录的确提供了STLF发展现状的一个合理的剖面，其中许多文献是最近出版的，正是它们补充了本文。

二、系统负荷

系统负荷是所有系统节点上逐个负荷的总和。原则上讲，如果每个负荷的模式知道的话，就可以决定系统负荷的模式。然而，单一负荷或用户的用电模式是完全随机的而且是难以预测的，在一个典型电力公司内部就有很多单个用户模式。这些因素使得不可能用外推单个用户的估计模式的方法来预测系统的用电水平。然而，好在是总负荷呈现出明显的可以用统计学的方法予以预测的用电模式。

系统负荷的行为受一系列因素的影响，我们把这些因素分为如下主要的四类：

1. 经济的；
2. 时间；
3. 气象；
4. 随机因素。

为了模拟系统负荷，必须弄清楚每一因素对用电模式的影响。下面我们扼要讨论这些因素的影响。

（一）经济因素

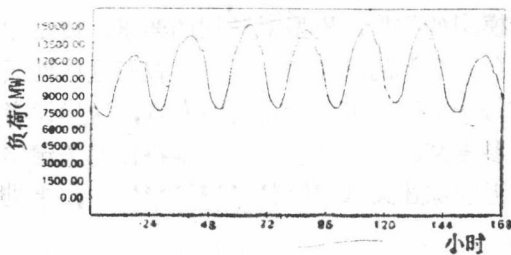
每个电力公司都在某一经济环境下运行其系统，而经济环境对负荷用电模式有明显的影晌，诸如供电区的人口、工业化水平、农事季节的变化、用电设备的特性和普及饱和程度、人工气候的发展和更一般的经济趋势等都给系统负荷的增加或减少以显著的影响。另外，公司初始制定的规划，如电价设计方面的变动和负荷管理规划也影响到负荷。一般这些经济因素在一个比一周要长的多的时间内起作用，重要的是在逐年地或逐季地修改预测模型的时候考虑这些因素。然而，经济因素不会明显地体现在短期负荷预测模型中，因为它们具有较长的时间尺度。

（二）时间因素

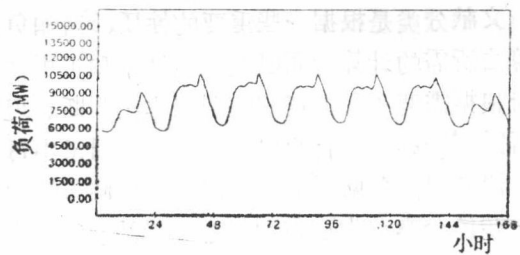
三个基本的时间因素——季节因素，周循环以及法定的和宗教性质的假日对负荷模式的影响起重要的角色，季节因素的变化决定一个电力公司是处于夏季或冬季高峰。随着季节性的变迁，诸如日照长短和温度升降，负荷的模式逐步产生变化。另一方面，这里有季节性的事件，它们使用户模式带来突发的但重要的结构性修改。这里有实行或取消夏时制时的转变，电价结构的改变（一天中不同时间或季节性负荷），学年的开始和假期（圣诞节—新年期间）生产性活动的显著减少。

负荷以一周为周期的变化是供电区人口的一个有次序的工作——休息模式。对典型的季节性周，这里有些定义好了的负荷模式，图3给出了一个具有夏高峰的电力公司的典型的夏季和冬季的周负荷模式的例子。

法定和宗教性假日的存在一般能显著地降低负荷数值至“正常”值之下。不但如此，在假期的前几天和后几天，要过一个“长周末”的趋势使我们能觉察到用电模式的变动。



典型的夏天周负荷曲线星期日~星期六



典型的冬天周负荷曲线星期日~星期六

图3 具有夏高峰电力公司的典型周负荷模式

（三）气象因素

气象条件对负荷模式有显著的影响，这是因为各电力公司都有很大的对天气很敏感的负荷分量，如取暖、空调和农业灌溉。

在许多系统中，温度是对负荷影响最大的气象变量之一。一天之中，温度对正常值的变化可能引起负荷足够大的变化，从而要对开停机计划作大的改动。而且前几天的气温也会对负荷曲线形状有影响，例如几天的连续高温天气可能创造一个新的系统峰值。对一个地理和气候不一致的系统，也许需要考虑几个地区的温度作为系统负荷的变量。湿度也是一个变量，它以与温度相似的方式影响系统负荷，特别是在高温和潮湿的地区。雷阵雨也会对负荷有很大的影响，因为它带来了气温的变化。其它影响负荷行为的气象因素还有风速、降水量、云层和照度。

（四）随机扰动

在这一节我们汇总了一组随机事件，它们引起负荷模式的偏移，但却不能由上面讨论的因素予以解释。一个电力系统始终处于随机的扰动之中。事实上，系统负荷就是由大量的各自不同的负荷组成。在大量的很小的扰动以外，还有大的负荷——钢厂、同步加速器、风洞，它们的运行能引起电力负荷大的波动。因为这些大设备何时投入运行电力系统的调度员是不知道的，所以它们代表了大的无法预测的扰动。还有一些事件，如大规模的罢工、工厂关闭和特别的电视节目，它们的发生是预先可知的，但它们对负荷的影响就是不确定的了。

三、文献分类

文末参考文献分类的目的是为研究STLF或欲对其概貌有所了解的读者提供方便。从文献目录中可见,有大量参考文献。然而,这里几乎不存在根本性的差别。由于问题的性质,从这些有价值的论文中,很难判断一种模型和技术比另一种更好。原因是大电力公司的电力负荷的性质,如上节所述,系统负荷是一个随机的不固定的过程,这个过程包含了成千上万个单独的分量,它们的行为不遵循任何已知的物理规则,是变化无常的。结果是,所有的宏观模型事实上都是经验性的,只能通过广泛的实验来客观评估。

我们的观点是,就负荷预测方法而言,最好的试验是看它在实际控制中心的环境下,在长达至少二年时间内的性能如何。到时候,我们才能评价其模型的能力:很好地完成季节的变化,正确地跟踪参数变量,有效地处理坏的或非正常的的数据,较好地与操作员进行人一机对话。不幸的是,如果我们排除了经典的基于操作员的负荷预测系统,那么只有为数不多的几样技术在实时操作系统环境下已经被实现,或者用实时数据被试验过。

文献分类是根据一些重要的性质,诸如负荷模型的类型,模型对数据的要求,模型和预测算法所需的计算时间以及实验结果和可用性。负荷预测的潜在用户不得不权衡这些性质,同时根据需求和可用资源的型式作出判断。现对每一类选择了一些合适的文章,这样读者就不再需要去翻阅所有的出版物。文献目录中每一参考文献都包含表征其基本特性的关键字。

读者同样会见到某些短期负荷预测方面的最近的综述文章^[3, 10, 23, 39, 48, 50],来进行进一步的分类和作这方面状态的描述。

下述STLF著作的分类是根据所使用的负荷模型的类型。某些重要的方面,诸如数据要求、计算时间和实验结果,在提到每一负荷模型类型时加以讨论。分类时考虑了两个基本的模型:峰值模型和负荷形状模型。峰值模型基本上具有单一型式。我们将负荷形状模型归结为两种基本的类型,每一种又具有它的子类型,即

1. 日时间*

a. 隐含时间函数模型的合成

b. 谱分解模型

2. 动态的

a. ARMA模型

b. 状态空间模型

我们将详细讨论每一种模型。

(一) 峰值负荷模型

这里,只有日或周峰值负荷被模拟成气象的函数,时间在这种模型中不担任角色。这种模型具有如下典型的形式:

$$\text{峰值负荷} = \text{基本负荷} + \text{气象因素分量} \quad (1)$$

$$\text{或者} \quad P = B + F(W) \quad (2)$$

式中基本负荷B是一个平均的不受气象影响的负荷分量,而气象因素分量 $F(W)$ 与B相加。气象变量W可以包括峰值负荷出现时的温度或者是预报的和历史温度的结合。湿度、照

*原文time of day,意为一天中的某一时刻,译为日时间,下同,——译者

度、风速和降雨量同样被考虑在这种模型之中。函数关系 $F(\cdot)$ 是根据经验计算的,它可能是线性的或非线性的。峰值负荷模型的例子可在文献^[6,58,63,69,70]中找到。

峰值负荷模型的优点是它的结构简单以及在初始化和修改时较低的数据要求。模型的参数通过线性或非线性的回归估计而得。这种模型的缺点是它们既不能定义峰值发生的时间也不能提供负荷曲线形状的任何信息。因为模型基本上是静态的,动态的现象,如跨周期的相关性无法预计。

(二) 负荷形状模型

这些模型把负荷描绘成预测时段内的一个离散的时间序列(过程)。负荷采样时间间隔一般是一小时或半小时,被测量的量一般是采样周期内的能量消耗,以兆瓦计。许多预计技术描述了负荷曲线的形状,它也包含了峰值负荷。然而,因为峰值负荷很难预测得准确,所以希望把负荷曲线模型和专门的峰值负荷模型结合起来^[6]。

基本上有两种型式的负荷形状模型:日时间的和动态的模型。把这两种基本型式结合起来是可能的。

日时间模型:日时间模型把预测时段 T 内的每一不连续的采样时间 t 的负荷 $Z(t)$ 定义成一个时间序列

$$\{Z(t), t=1, 2, \dots, T\} \quad (3)$$

这是一个最简单的形式,日时间模型根据以往观察的负荷行为存贮了 T 个负荷数据。某些电力公司仍然使用前一周的实际的负荷模式作为一个模型来预测本周的负荷。还存贮了一组一年中若干典型周或典型气象条件的曲线,如潮湿的、干燥的、多云的或刮风日子的,并把它们与最近的周负荷模式有机地结合起来作出预测。在这种情况下,操作员的判断决定了最后的负荷预测量,而不采用明显的数学公式去描述模型的机理。这也许是专家系统应用的一个潜在领域。专家系统将积累操作员所遵守的规则^[2]。有关这种启发式模拟方法的文献不多,一些有关的基于群集分析和模式识别的研究工作可在文献^[18,25,28,52]中找到。

一种更普遍的日时间模型取如下形式:

$$Z(t) = \sum_{i=1}^N \alpha_i f_i(t) + V(t), \quad t \in \tau \quad (4)$$

式中,时刻 t 的负荷 $Z(t)$ 被认为是有限个显式时间函数 $f_i(t)$ 之和,这一时间函数在24或168小时的范围内一般是正弦的,与预测提前的时间有关。 α_i 被作为随时间缓慢变化的常数, $V(t)$ 表示模型误差,假设是白噪声。这种模型被认为在包括过去、现在和将来的一个时段 τ 内是有效的,将来的时间包括最大的提前时间。

当 $f_i(\cdot)$ 是被事先选定的显式时间函数时,如正弦的,参数, α_i 通过对一组过去的负荷观察值 $\{Z(t), t \in \tau_{\text{past}}\}$ 简单的线性回归或指数平滑分析来加以估计, τ_{past} 是一个从现在起往前推的一个时段^[59]。这种模型的例子可在文献^[17,46,49,51,60,64]中找到。这种模型的优点是结构简单,同时模型参数可以通过线性回归指数平滑非常简单地予以修改。这些方案的本质是要求较少计算时间的递归算法,可以被用来修改参数,当有新的负荷数被测量时,还可以来修改预计本身。相反,日时间模型不能精确地代表负荷过程的随机的相关特性或它与气象变化的关系。结果是,当气象模式在迅速地变化时,除非对未来很短的时段,否

则系数 α_i 就不适用了。从而, 这将给较长时间的预测带来精度的问题。

还有一种第二类的日时间模型, 它们建立在谱分解的基础之上。模型具有式 (4) 的基本形式, 然而, 这里时间函数 $f(\cdot)$ 代表特征函数, 它相当于负荷时间序列的(在滤去趋势和周期性以后)的自相关函数。这种方法是基于 Karhunen—Loeve 谱分解的扩充^[43, 71]。它的优点是被选来代表负荷时间序列的时间函数是最优的, 因为他们能比较近似于它的自相关函数, 即它的二阶概率行为。从而, 这一方法中时间函数的合成能够代表平稳的有色随机负荷; 其精度比任意选取的时间函数要高。尽管系数 α_i 是利用线性回归技术估计的, 特征函数的识别要求过程自相关矩阵的一个近似解和相应特征值问题的解。这种识别的方法对实时回归算法是不适用的, 因为它需要更多的计算时间。如果过程是几乎平稳的, 则识别部份只是在不常用的时间段才是需要的。这种技术对突然的和大的气象变化条件下的误差是敏感的, 因为这些效应并未明确地被模拟。尽管谱分解模型在理论上比起其它的日时间模型更可靠, 但是它的实际的优越性并未明显地被证实, 结果是只有少数几家公司看起来相信这种方法^[12, 42, 68]。

动态模型: 动态负荷模型考虑到负荷不但是时间, 而且也是现时行为的函数, 以及气象和随机输入的函数。动态模型也有两种基本形式, 自回归移动平均或 ARMA 模型和状态空间模型。

ARMA模型: ARMA模型取如下通用形式:

$$Z(t) = Y_F(t) + Y(t) \quad (5)$$

式中, $Y_F(t)$ 是一个首先决定于日时间和某天正常气象模式的分量, 它能用一个由式 (4) 给出的周期时间函数表示。 $Y(t)$ 是一个附加负荷残差项, 它描述气象模式与正常的偏差和随机相关效应的影响。把残差负荷作为附加的, 这是因为它的作用与日时间分量相比是很少的。也有用非线性模型去描述周期和残差分量的相互作用的, 但不常见^[11]。残差项 $Y(t)$ 可以用如下形式的 ARMA 过程来模拟, 即

$$Y(t) = \sum_{i=1}^n a_i Y(t-i) + \sum_{k=1}^{n_u} \sum_{j_k=0}^{m_k} b_{jk} u_k(t-j_k) + \sum_{h=1}^H C_h W(t-h) \quad (6)$$

式中, $u_k(t)$, $K=1, 2, \dots, n_u$ 代表 n_u 个气象因素输入。气象变量的作用被认为是明显的。这些输入是某日某时的温度、湿度、光照和降水量等量与正常水平偏差值的函数。输入 $u_k(t)$ 也可以代表系统内不同地区所测量的气象效应的偏差。过程 $W(t)$ 是一个零均值的白噪声过程, 它代表不定的因素和随机的负荷行为。参数 a_i , b_{jk} 和 C_h 以及模型阶参数 n , n_u , m_k 和 H 都是假设为不变的, 但却是未知的参数, 有待将模拟的模型数据与观察的负荷和气象数据拟合而得。

文献中给出了与由式 (5) 和式 (6) 描述的基本模型不同的许多模型。名称也各异, 有 Box—Jenkins, 时间序列, 传输函数, 随机模型, ARMA 和 ARIMA, 因为这些模型之间差别不大, 我们宁愿采用一个统一的方法, 并把注意力集中到这些模型的共性上, 因此, 我们把这些模型统称为 ARMA 模型。关于 ARMA 模型的详细说明, 请读者参考教科书^[66]。

有些作者^[13, 14, 44, 58]已采用了显式表达周期负荷分量的方法, 如式 (5), 而另一些作者^[8, 16, 19, 21, 27, 30, 38, 47, 60]先将负荷数据进行预滤波, 以消除周期分量作为显式的

时间序列。预滤波基本上是用定义如下一个新的负荷过程的方法来实现的。

$$Z'(t) = Z(t) - Z(t - t_p) \quad (7)$$

式中, t_p 是日时间分量的周期(一般为 24 或 168 小时)。所以,产生的过程 $Z'(t)$ 是与周期项无关的,同时满足类似于式(6)的 ARMA 方程式。这样做的好处是较规范的技术可以用来识别得出的 ARMA 模型的参数^[37, 60, 66]。预滤波的不足之处是这种方法基本上相当于对一个几乎肯定包含测量和模拟误差的过程进行微分。结果是放大了测量的误差,导致相应的模拟不准确。显式的日时间分量模拟不需要预滤波,因而不受这种类型的预滤波误差的支配。然而,对于这样的模型,必须使用非线性的参数估计方法去识别模型参数。这将使参数识别阶段的计算工作量稍有增加。负荷模型中常数偏移或时变趋势的存在也能被合适的预滤波^[19]或它在时间分量中的时间多项式的显示表达或予以处理。

只有某些 ARMA 模型把气象因素作为一个输入(请参考文献目录中带关键字 ARMA 和 W 的文献)。那些不包括气象因素的模型自动地修改某些参数,去考虑气象的变化对负荷的影响。这种方法在气象条件迅速变化期间是不能满足的,在这种条件下不再满足负荷过程是平稳的条件。在许多最近的 STLF 模型中,气象因素是被明显地考虑进去的。一些可用的 ARMA 模型通过附加的显式输入来描述气象的影响,如式(6)^[8, 14, 18, 21, 56],而其它的模型依赖于更探索性的方法,在应用 ARMA 模型于被校正的负荷之前,负荷过程先被纠正温度的影响^[10]。最重要的气象因素输入是根据温度的偏移,它总是被表达为实际和“正常”温度差的非线性函数。这种函数考虑了在不同季节温度对负荷的不同影响,死区和其它的非线性特性。将实际气象的影响与 ARMA 输入相联系的模型基本上由实验而得,并随系统而异(请阅本段引用的文献)。然而,有些非线性的影响是众所周知的。在夏天,随着温度的升高,大多数系统的负荷是增加的。在冬天,这种现象就倒过来。有理由假设,不是气象变量的绝对数值,而是它在一天中的特定时刻和一年中的特定天中与“正常”水平的偏差影响着负荷。日时间或周期分量将考虑气象对电力用户长期的季节性的影响。

ARMA 模型的参数识别比起日时间模型来一般需要更多的计算时间,然而,这种额外的计算量对得到更加健全的,包括动态的、气象的随机效应的模型是必不可少的。长期的使用表明,需要调整的参数不多,却得到较好的预测性能。在任何情况下,因为参数识别是不经常进行的(一天一次),所以对所讨论的方法来说,计算工作量对 EMS 计算机不是一个主要的因素。对一个一般的 ARMA 模型的参数识别能用包括求解 Yule—Walker 方程的递归方法取得^[8, 60],或者利用最大似然方法^[44],这基本上是一个非线性回归的算法。ARMA 模型某些形式的可调系数通过线性回归技术予以识别。这些模型是 $Y(t)$ 中的 AR (Auto Regressive—自回归) 和 $u_k(t)$ 中的 MA (Moving Average—移动平均),但随机输入 $W(t)$ 中却不是 MA。那些显式描述日时间分量的模型一般要求应用非线性回归方法去识别动态模型和周期分量的参数^[56]。把模型的 AR 部件略去,就可以避免应用非线性回归^[14],但却失去模拟负荷短期随机相关性的能力。读者可以阅读许多详细描述上述参数识别技术的优秀的文献^[22, 50, 59, 65]。

一般说来,模型参数的修改不是一个计算量很大的任务,即使要求非线性估计问题的迭代求解时也是如此。在参数能用线性回归而加以估计的模型中,当获得新的负荷值和气象数据时,参数的修改能用在在线递归的方法实现。这时频繁的参数修改是不必要的,除非模型非常简单,如纯日时间模型,它才需要不断地修改。对比较复杂的模型,如 ARMA,模型的结构

和参数在几天之中保持不变。在以小时计的时段内,这样的参数修改事实上是不希望的,特别是在非正常负荷行为期间,模型参数肯定是不修改的。对于 ARMA 模型,按日修改参数可能是足够了。在这种情况下,前24小时的数据在“消除”了它们的非正常行为后,被加到数据集中,而老的24小时数据被移去。日参数的修改不是什么紧急的任务,它可以在计算机最不忙的时候去做。

状态空间模型:众所周知,一个ARMA模型可以被转换成状态空间模型或相反^[22],所以在概念上,两种模型之间并不存在根本的差别。然而,在文献中提出了许多状态空间模型,这些模型增加了在典型的ARMA模型中不常出现的一些结构。在这些模型中,t时刻的负荷 $Z(t)$ 一般由下式所得

$$Z(t) = C^T X(t) \quad (8)$$

$$\text{式中 } X(t+1) = AX(t) + Bu(t) + W(t) \quad (9)$$

时刻t的状态向量由 $X(t)$ 表示,气象变量的基输入向量是 $u(t)$,随机白噪声向量是 $W(t)$ 。矩阵A、B和向量C认为是常量。有许多从这种基本模型变化而来的模型,有时候,状态 $X_i(t)$ 、 $i=1,2,\dots,N_s$ 可以代表一周内某一天某一小时的周期的负荷分量,或这模型的一个参数,或负荷与气象相关输入的组合。状态空间和ARMA模型之间的一个区别在于,可用的状态空间模型技术假设,定义负荷周期分量的参数是随机过程。在本质上,这就允许我们使用先前的有关它们数值的某些信息(实际上合理的假定)。这将 对用 Bayesian 技术进行参数估计方面是有帮助的。然而,这样一个先前参数的信息在ARMA模型中也可以加以利用。在某些已提出的状态空间模型中,矩阵A和B是非常稀疏的,且是已知的^[4,28,55,57],而另一些状态空间模型则要求对整个A矩阵进行识别^[30,61]。比起ARMA模型来,状态空间模型的优点目前还不是很明显的,需要更多的实验比较。证明状态空间方法优越性的一个可能的领域是发展母线负荷预测,这里母线负荷显示出高的相关性。

(三)STLF发展水平小结

在两种主要的STLF模型——峰值模型和负荷曲线模型中,后者比较通用。负荷曲线模型有较大的适应性,且一般更准确。纯日时间模型至今已几乎被动态模型所取代,因为日时间模型不能精确地代表与时间相关的随机效应和气象影响。两种主要动态模型,ARMA和状态空间模型,利用随机和气象输入。从已出版的文献判断,看起来ARMA比状态空间法更通用。也许是因为前者要求较少的说明变量和参数。计算时间随可用的STLF技术不同而异,但不管在什么情况下,它不是主要要考虑的。因为离线和在线的计算时间都是适中的。STLF文献中主要缺少的部分是利用实际数据,特别是在线环境下的经验报告。文献中也缺少各种STLF方法应用于一个基准系统的标准数据组的性能比较的研究。

四、实际的考虑因素

本节通过某些为开发一个STLF模型和程序所需要的主要步骤引导读者考虑一些实际的约束和要求。尽可能地引用参考资料以提供辅助的信息和经验。我们将特别讨论模型的形成和选择、预测的算法、性能估计和实现等实际方面。

一个一般的负荷模型和预测程序以如下的步骤应用于STLF问题:

1.模型的形成和选择;

2. 模型参数的识别和修改;
3. 试验模型性能和修改预测;
4. 如果性能不满足要求, 返回 1 或 2, 否则返回 3。

模型的性能必须不断地给予监视, 一旦一个合理的模型结构已经被建立起来了, 模型性能的变坏应该首先通过步骤 2 用很好地调整模型参数的方法予以纠正。当合适的模型已选择好了, 不必经常地去进行模型结构的改变。每一小时或半小时修改一次模型状态和负荷预计。

伴随短期负荷预测的计算机的要求是适度的。消耗在 EMS 应用软件的预测部分上的时间片是很小的, 这是因为预测程序在计算时间上不是很多的, 执行的次数也不多。适当的磁盘存贮器必须提供来用于历史负荷和气象数据的存贮, 这些数据是作为预测模型的初始化和修改用的。

调度员乐于应用那些使用方便和特别是在关键的时候工作可靠的预测软件包, 运行人员比较关心高峰期间而不是非高峰期间或假期的预测结果。模型必须在这些关键的时刻工作准确和可靠。

(一) 模型的形成和选择

在选择 STLF 模型时首先考虑的是预测值的性质, 所需预测提前时间和预测的应用。也许需要多个模型来预测系统日高峰负荷、一天中特定时间的系统负荷和 (或) 周系统电量。在某些情况下, 可有几个模型来预测在统计上相结合的同样一些量^[61]。多预测模型的应用提供了一个有效的检查系统, 增加了预测结果的可靠性。

对一个考虑中的特定模型, 首先进行普遍意义上的测试。回答以下几个基本问题:

1. 模式的意义?
2. 是否所有影响特定系统负荷的因素显式地或隐式地被考虑了?
3. 模型具有物理意义吗?

以上这些问题必须在进一步工作之前得到肯定的回答。

在形成和 (或) 选择合适的模型时一个重要的考虑是模型的经济性, 基本的争论点是:

1. 独立的或说明变量的数目;
2. 预测的方便性和每一说明变量相应的不确定性;
3. 可调参数的数目。

一般说来, 较少说明变量和可调参数的模型是较好的。这样的模型便于初始化、修改、调整和运行。

模型的形成和选择的进一步考虑是其对数据的要求, 不同模型对数据的要求与模型的性质有很大的关系, 一般地, 要求大量初始数据的模型有:

1. 非线性模型;
2. 涉及到移动平均项的随机输入模型;
3. 带气象说明的模型;
4. 带许多参数的模型。

相比之下, 其系数呈线性的模型, 如日时间模型, 在初始化和修改时一般要求较短的数据周期。

对数据的要求存在着两重性, 一方面希望在相关的和独立的或说明变量之间建立一个尽可能长久的关系, 这必然要求一组复盖一个长时期的历史数据, 另一方面, 还要使模型能够灵活地去反映基本过程的任何变化, 这就不得不要求一组仅仅复盖最近时期的较短期的数

据。这样就排除了任何以前的过程，它们现在已不再有效了。这一考虑对于在特殊情况下把节能作用的模型和在一般的情况下把新的管理要求程序并入模型中也许是特别恰当的。

高数据密度的模型不利于应用和修改，一般地，要求数据较少的模型较好。譬如说，如果模型的初始化要求几个月的数据，人们也许会问假定季节性变化和不正常性在这段时间内可能是显著的，这个模型是否能合理地代表负荷。三至六周的数据是恰当的。另外，还要注意数据库也必须包含关于一年中特定日子的信息，这些日子是以年为周期的。

某些判断显然涉及到某一特定系统负荷模型的形成。得知了负荷拟型和预测方面的现状之后，下一步就是合理地去发展一种模型，使之不但能描绘负荷曲线的形状，而且还有其动态的、气象的、日时间的和随机的效应。那些只能反映峰负荷或者不能显式地模拟气象效应的模型，尽管便于建立和修改，但不能提供较通用模型所具有的精确性和灵活性。在模型形成阶段，人们也许会把对模型的选择局限于那些根据数据类型和可用计算设备最适合于用户需要的模型。然而，我们必须保持开宽的视野，并对几个类型的模型进行试验，因为没有结论性的证据可以指明这个模型比其他更优越、同时也必须注意到许多模型可以被识别并在可接受的计算时间和数据要求之下运行，所以这判断可能不是那末严格。最终的判断是模型对真实数据的预测性能，因为没有实验是很难预言的。

STLF模型的初始化工作要求检查包含至少二至三年的小时负荷值和气象数据的数据库。尽管模型的参数在某一范围内被跟踪几个季度，但年负荷行为中有许多特殊的日子或不连续的日子，这些日子虽然一年只发生一次，然而必须把它们作为一个日时间分量的一个专门项（假日、夏时制的实行、学校开学和放假）予以识别和模拟。另外，在进行识别之前，对那些非正常行为的负荷加以检验，这些非正常行为是由诸如罢工、停电、选举日或特别的电视节目引起的。这些非正常行为必须加以识别，且把它们排除在“清洁”的初始化数据之外。在这一阶段，有经验的负荷预测操作员的输入是基本的。在初始化阶段，根据先前的经验或简单的校正试验，对气象输入的要求也可以被设置。

（二）预测的算法

预测的算法是与形成的负荷模型的类型紧紧地联系在一起的。一旦选择好了负荷模型，预测的算法也就基本上被确定了。所有预测的次序都遵循以下步骤：

1. 把从模型初始化阶段或从参数识别 / 修改算法得到的模型参数估计值代入负荷模型中；

2. 确定预测提前时间；

3. 如果气象变量参与模型之中，输入其预报值和误差估计值（如有的话）；

4. 如果模型是动态的，利用递归的线性估计方法，估计现在系统的状态（动态方程(6)的初始条件）^(22, 59)；

5. 利用模型计算被预测的负荷，估计参数，确定的提前时间，如果模型要求的话，还有预报的气象变量和实际的状态估计值以作为初始条件，如果模型有白噪声随机过程输入，为了预测起见，它是由其均值估计的。

6. 如果模型允许的话（动态随机模型）计算误差的方差。

这样，在日时间或非动态模型中，负荷预计就是很简单的事了，把提前预测时间和相应的气象变量代入到被估计系数参数化的负荷模型方程中即可。而另一方面，动态模型包含了差分方程，这要求初始条件估计（状态估计）去启动前向模拟，和一个将来输入的估计，也就是气象预测，进行时间上的前向模拟。状态是由递归线性估计过程来估计的^(22, 59)。这一过

程利用最近的负荷和气象变量的数值来修改状态。因为差分方程是递归的,将来某一时刻的负荷值只有通过计算从现在起到那一时刻间的所有负荷预测值才能得到。动态模型的这一预测要求比起非动态模型增加了在线计算工作量,然而,这一额外的工作量是在现代控制中心计算机的能力之内的。

(三) 参数识别

在应用参数识别技术去“净化”数据库之前,不但要计及季节性负荷变化,同时还要考虑一年至下一年的负荷增长/减少的趋势。处理这一时变的方法是如式(7)所示的以年为周期的数据前置滤波^[16,18],从前置滤波负荷过程中消去季节性变化,从而被认为是平稳的。第二种更加通用的处理季节性变化的方法是假定,在季节中负荷模型是随时间缓慢变化的。数据的移动时间窗口被用来识别模型参数,这些参数在移动窗口内或在将来的预测时段内被认为是不变的。这些时间窗口长达三至六周,依其处在一年中的什么时间而定。为了增加用以识别的可用数据量,上一年的同一时段的移动窗口可结合到大的数据库中去。

在上一节模型分类中可见有关在参数识别步骤中使用各种方法的讨论和例举的许多有关的文献。

(四) 性能估计

一个给定的短期负荷预测系统的性质可以用以下各项予以估计:

1. 模型的准确度;
2. 应用程序使用的方便性;
3. 坏/非正常数据的检测和纠正能力。

准确度:模型准确性的评价要求预计误差,也就是在预测周期的每一时间点确定的负荷预测值和负荷“测量”(实际)值之间的差异。一般是用统计的方法根据预测误差的标准偏差来衡量模型的准确度。其它准确度的衡量方法也是可能的,如用最大误差或加权二乘法,这里权值的选取是对高峰时采用最大的权值,而非高峰时采用较小的权值,但这些方法不常使用。事实上,很难找到这样一些STLF系统,它们24小时预测的均方根误差小于高峰负荷值的2~3%。这也许构成了一个模型拟合性能好坏在统计上的极限,总的来说,它代表了固有的负荷噪声分量。然而必须注意,实际上的24小时预测误差与负荷的类型很有关系,它是生活的、工业的、商业的分量,地理的位置和季节变化的组合体。

在更一般的意义上,两个主要的因素——预测提前时间和说明变量的不确定性限制了预测模型的准确度。预测提前时间越长,预测的准确度越差。模型中说明变量数越多,模型的不确定性也越大。当说明变量本身就具有很大的不确定性时,更是如此。而且,必须认识到不同的预测气象变量有不同的预测误差。譬如说,预测气温要比预测降雨量容易的多。应该避免那些难以预测的独立变量,以便消除产生大的固有误差预测的可能性。

两个或更多个不同短期负荷预测模型准确度的比较应该在尽可能相似的实际的运行条件下进行。为使比较工作有意义,应在足够长的周期的许多子周期内利用变量预测值进行评价。这种方法使得不同模型的性能在一个大的数据范围内均匀地进行比较。

模型性能试验由以下方法系统地完成,即证明一步预测误差 $e(t)$ 形成一个白噪声过程,这里

$$e(t) = Z(t) - \hat{Z}(t/t-1) \quad (10)$$

式中, $Z(t)$ 是 t 时刻的负荷,变量 $\hat{Z}(t/t-1)$ 是给出到 $t-1$ 时刻的实测值和气象数据后,