

电网物资需求预测中区间估计的应用研究

胡永焕¹, 李俊颖¹, 姜 晨², 漆 璇³, 董凤娜⁴

¹国网上海市电力公司, 上海

²国网上海市电力公司崇明供电公司, 上海

³国网上海市电力公司青浦供电公司, 上海

⁴上海久隆企业管理咨询有限公司, 上海

收稿日期: 2024年10月29日; 录用日期: 2024年11月18日; 发布日期: 2024年12月31日

摘 要

在电网工程物资管理中, 需求预测的有效性十分重要。有效的需求预测可以防止采购过剩和供应不足。统计分析是常用的需求预测方法, 其中尤以回归分析应用较为成熟。简单回归分析仅给出一个结果数值, 而不反映预测的可靠性。对于电网物资供应链决策而言, 由于资金和库容条件限制, 物资采购总量只能维持在安全区间。管理者在面临需求不确定性时, 需要明确当前采购量的可靠性。本文基于区间估计建立回归预测模型, 并给出一些应用案例。预测模型同时返回需求区间和可靠性, 这样需求不确定性被量化为风险概率, 为供应链管理提供二元决策模型。二元决策模型的优点在于管理者可以根据可靠性目标分配电网物资。基于显著性检验方法, 二元决策模型也可以衡量某一段时间内协议库存满足率。

关键词

电网, 需求预测, 区间估计, 可靠性

Research on Interval Estimate Application of PGM Demand Forecasting

Yonghuan Hu¹, Junying Li¹, Chen Jiang², Xuan Qi³, Fengna Dong⁴

¹State Grid Shanghai Electric Power Company, Shanghai

²Chongming Power Supply Company of State Grid Shanghai Electric Power Company, Shanghai

³State Grid Shanghai Electric Power Company Qingpu Power Supply Company, Shanghai

⁴Shanghai Jiulong Enterprise Management Consulting Co. Ltd., Shanghai

Received: Oct. 29th, 2024; accepted: Nov. 18th, 2024; published: Dec. 31st, 2024

Abstract

The effectiveness of demand forecasting is crucial for power grid material (PGM) management. It

文章引用: 胡永焕, 李俊颖, 姜晨, 漆璇, 董凤娜. 电网物资需求预测中区间估计的应用研究[J]. 现代管理, 2024, 14(12): 3237-3243. DOI: 10.12677/mm.2024.1412389

contributes to both surplus purchase and short supply prevention. Statistical analysis is the most common method for demand forecasting. Especially, regression is rather apply proven. Simple regression analysis only returns amount as a result which does not reflect reliability of demand forecasting. In terms of PGM supply chain decision, due to the limitation of budget and capacity the purchase amount must remain on the safe side. The managers need clarify the reliability of current amount when they face the uncertainty of demand. This paper establishes a regression forecasting model based on interval estimate and gives a few application cases. Forecasting model simultaneously returns the interval and reliability about demand forecasting. So that the uncertainty of demand being measured as risk probability affords binary model for supply chain management. The advantage of binary model is that managers could distribute PGM according to the reliability target. Binary model also could measure the probability of agreed stock fulfillment.

Keywords

Power Grid, Demand Forecasting, Interval Estimate, Reliability

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

物资需求预测是指项目正式提报需求前,通过各种技术手段,合理估计出所需物资品类和数量。电网项目因具有公共性质,按法律规定必须进行招标采购。考虑到招标成本和规模效应,电网物资供应链管理中,高频物资通常采取协议库存采购方式。协议库存采购是将固定周期(比如一年)的物资需求,一次性集中采购到位,待到项目需求明确后再行配送。协议库存长周期采购方式,意味着供应链不可能遇缺即补、随买随用。因此需求预测的有效性,对供应链韧性就十分重要[1]。若需求预测数量偏高,供应链将出现采购过剩、资源挤占、甚至履约超期问题。反之,若需求预测数量偏低,供应链将面临断供风险,影响项目正常进度。

2. 理论方法研究

目前,电网工程需求预测的主流方法有线性回归法、指数平滑法、时间序列法、灰色预测法、神经网络法和马尔可夫链法[2]。指数平滑法和时间序列法适用于具有时间趋势的年度总需求预测,无法用于单个工程预测。灰色预测法仅适用于中短期指数模型,通用性过窄。马尔可夫链无法基于工程特征建立模型,分析预测精度不可控。相比之下,线性回归法和神经网络法可以基于工程特征作预测,神经网络算法对抗小样本极端值的效果不佳,而线性回归法则不受该问题限制[3] [4]。

电网企业在回归分析建模的实践中,总结出了不少技术创新和成功案例[5] [6]。简单的回归分析只返回一个结果数值,建模者对预测准确率的提升手法极为有限。无论如何改进算法和筛选样本,回归分析都无法做到完全准确。事实上,由于电网发展建设固有的不可预见性,以100%准确性为目标的预测理念并不可行。受限于资金和库容条件,供应链管理实务中,通常选择将物资采购量维持在一个安全水平,让风险事件的发生概率小于可接受的水平[7] [8]。本文从实务需求出发,设计了一套“数量-概率”二元决策模型。二元决策模型是应用回归分析中的区间估计,将需求的不确定性被量化为风险概率,让管理者可以明确当前采购量的可靠性。二元决策模型的优点,除了可以根据可靠性目标分配电网物资,实现“概率到数量”的正向分析,也可以衡量某一段时间内协议库存满足率,实现“数量到概率”的反向分析。

3. 回归分析的区间估计

在 n 个项目构成的分析样本中, 某种物资的实际需求量为 $Y = \{y_1, y_2, \dots, y_n\}$, 需求量的影响因素有 m 种, 取值为矩阵 $X_{ij} = (x_{ij})$, $1 \leq i \leq n, 1 \leq j \leq m$, 其中 x_{ij} 为第 n 个项目的第 m 个因素值[9]。简单回归分析, 建立线性方程 $Y = \beta X + \varepsilon$, 待定系数 $\beta = (\beta_0, \beta_1, \dots, \beta_m)^T$ 。求解模型待定系数 $\beta = (X^T X)^{-1} X^T Y$, 这样就得到了初步的回归公式[10][11]。回归公式仅仅求解了最逼近样本点的直线, 但样本可能原本就没有线性关系, 所以需要对回归公式(线性系数)作假设检验, 一般采取 t 检验法[12]:

$$1) \text{ 标准差的无偏估计 } s = \sqrt{\frac{1}{n-m-1} Y^T (I - X(X^T X)^{-1} X^T) Y}$$

$$2) \text{ 第 } i \text{ 种因素的检验统计量 } t = \frac{\beta_i}{s \sqrt{\left((X^T X)^{-1} \right)_{(i+1, i+1)}}} = \frac{N(0,1)}{\sqrt{\chi_{(n-m-1)}^2 / (n-m-1)}} \sim t_{n-m-1}, \text{ 其中 } N(0,1) \text{ 是标准}$$

正态分布, $\chi_{(n-m-1)}^2$ 是自由度为 $n-m-1$ 的卡方分布, t_{n-m-1} 是自由度为 $n-m-1$ 的学生氏分布。

3) 设置假设检验显著性水平 α , 该值表明第 i 种因素有多大概率与需求量是线性无关的。计算统计量 t 值的概率 $t_{n-m-1} \left(\frac{\alpha}{2} \right)$, 再与显著性值作比较, 若 t 概率低于显著性, 则说明第 i 种因素与需求量有线性关系。

4) 没有通过假设检验的因素, 应该被剔除并重新进行回归分析。

求解得到回归公式后, 就可以利用其对新增项目物资需求量进行区间估计了。区间估计有两种类型, 分别为置信区间和预测区间[13]。置信区间是根据新增项目的影响因素值 X_{n+1} , 估计 Y 平均值在置信水平 α 下的区间, 估算公式为 $X_{n+1}^T \beta \pm s \sqrt{x_{n+1}^T (X^T X)^{-1} x_{n+1}} \cdot t_{n-m-1} \left(\frac{\alpha}{2} \right)$ 。预测区间则是根据影响因素值 X_{n+1} , 估计 Y 个别值的区间, 估算公式为 $X_{n+1}^T \beta \pm s \sqrt{x_{n+1}^T (X^T X)^{-1} x_{n+1} + 1} \cdot t_{n-m-1} \left(\frac{\alpha}{2} \right)$ 。显然给出个别值的预测区间, 比给出平均值的置信区间更宽[14]。

电网物资需求预测的应用中, 协议库存剩余份额只要高于预测区间上限, 即可视为满足项目未来开工需求。这样假定区间估计的置信水平为 α , 那么供应可靠性为 $1 - \frac{\alpha}{2}$ (或作单尾概率检验)。若已知每个待建项目的影响因素, 可以根据可靠性水平, 求解每个项目需求均值的置信区间, 再进行上下限叠加求出总量区间。若仅知晓待建项目的总体影响因素, 则可以直接求解需求量的预测区间。

4. 高压开关柜和小车需求预测模型

需求预测对象为 10kV 开关站的高压开关柜及小车。开关柜按功能, 分为进线柜、馈线柜、分段柜、隔离柜和母线柜, 共 5 种。配套小车有验电小车、接地小车和检修小车, 共 3 种。需求预测以单个开关站作为样本, 选取近年来 40 项新建建设工程。首先, 从电力工程设计的基本原理出发, 选取高压进线数和出线数作为高压开关柜的需求影响因素, 建立线性回归模型, 分析结果如表 1 所示:

结果显示, 五种开关柜都通过了回归分析的显著性检验, 仅有隔离柜的拟合度不满 100%, 但仍然达到了 87.6% 的极高水平, 说明回归模型的预测效果十分优秀。同样地, 对三种小车作回归分析, 需求影响因素为高压开关柜总数, 结果如表 2 所示:

结果显示, 三种小车的需求量中包含常数项, 即单个开关站的小车需求可拆解为固定值和可变值两部分。此外, 拟合度 R^2 只有 50% 左右, 表明小车需求量中有一半比例的非线性因素(可测性偏低)。供应链管理者在保持供应稳定时, 需要设置一个更加高的可靠性。

Table 1. Regression analysis of HVSC
表 1. 高压开关柜回归分析

自变量	因变量				
	进线柜 y_1	馈线柜 y_2	分段柜 y_3	隔离柜 y_4	母线柜 y_5
斯皮尔曼相关性					
高压进线数 β_1	1.000	0.907	1.000	0.961	1.000
高压出线数 β_2	0.907	1.000	0.907	0.920	0.907
线性回归系数					
高压进线数 β_1	1.000	0.000	0.500	0.491	1.000
高压出线数 β_2	0.000	1.000	0.000	0.006	0.000
常数 β_0	0.000	0.000	0.000	0.041	0.000
t 检验显著性					
高压进线数 β_1	0.000	0.000	0.000	0.005	0.000
高压出线数 β_2	0.000	0.000	0.000	0.841	0.000
常数 β_0	0.000	0.000	0.000	0.658	0.000
修正的拟合度 R^2					
-	1.000	1.000	1.000	0.876	1.000

备注：t 检验显著性为 0.05。

Table 2. Regression analysis of trolley
表 2. 小车回归分析

自变量	因变量		
	验电小车 z_1	接地小车 z_2	检修小车 z_3
斯皮尔曼相关性			
高压开关柜总数 β_1	0.693	0.755	0.713
线性回归系数			
高压开关柜总数 β_1	0.042	0.044	0.094
常数 β_0	1.172	1.298	1.656
t 检验显著性*			
高压开关柜总数 β_1	0.001	0.003	0.001
常数 β_0	0.005	0.006	0.038
修正的拟合度 R^2			
-	0.472	0.416	0.515

备注：t 检验显著性为 0.05。

获取未来一个周期内，待建开关站项目的高压进线数和出线数，估计各类高压开关柜的需求总量，其中供应链可靠性目标被设定在 97.5% (对应置信水平 5%)，解得开关柜和小车的需求量区间。同时，根据小车的协议库存实际剩余量，推算出无补充采购情况下，协议库存满足未来项目需求的可靠性。结果表明，检修小车供应可靠性偏低，断供风险概率估计在 16%，详细数据如表 3 所示：

Table 3. Demand forecasting model of HSVC and trolley
表 3. 高压开关柜及小车需求预测模型

预测区间	进线柜 y_1	馈线柜 y_2	分段柜 y_3	隔离柜 y_4	母线柜 y_5
下限	72	444	36	18	72
上限	72	444	36	56	72
置信区间	验电小车 z_1	接地小车 z_2		检修小车 z_3	
下限	14	13		33	
上限	45	48		96	
协议库存剩余量	55	75		80	
供应可靠性	99.8%	99.9%		83.9%	

备注：供应可靠性取单尾概率。

5. 架空导线和杆塔需求预测模型

预测对象为 10 kV 配网架空线路的导线和杆塔。导线规格覆盖 10 kV 和 0.4 kV 下所有截面铝线，共 7 种。杆塔分为钢管杆和锥形水泥杆，钢管杆以项目总重量计算，不区分具体杆数及杆型。需求预测选取近年来建设的 34 项线路工程，且必须来自上海不同地区，以保证设计档距和杆型消除区域性。架空导线的需求影响因素，定为 10 kV 和 0.4 kV 线路的设计总长，计量单位公里，分析结果如表 4 所示：

Table 4. Regression analysis of overhead line
表 4. 架空导线回归分析

因变量 (截面积)	自变量(线路总长)			
	斯皮尔曼相关性	线性回归系数	t 检验显著性	修正拟合度 R^2
10 kV 导线				
150 mm ²	0.812	0.317	0.000	0.917
240 mm ²	0.700	0.319	0.000	0.949
0.4 kV 导线				
25 mm ²	0.865	0.237	0.000	0.925
35 mm ²	0.884	0.236	0.000	0.940
70 mm ²	0.620	0.237	0.000	0.889
120 mm ²	0.940	0.244	0.000	0.997
185 mm ²	0.683	0.236	0.000	0.903

备注：常数项均未通过假设检验，本表已忽略。

回归分析结果与理论一样，架空导线需求量与线路设计长度呈高度线性相关。同理，将杆塔需求量代入模型中，以线路设计长度作为影响因素，钢管杆计量单位吨。分析结果表明，杆长规格在 15 m⁺的锥形水泥杆以及钢管杆，与敷设线路的长度没有线性关系。低压线路常用的 13 m⁻锥形水泥杆，与低压线路长度具有线性关系，分析数据如表 5 所示：

Table 5. Regression analysis of poles and towers
表 5. 杆塔类回归分析

因变量	自变量(线路总长)			
	斯皮尔曼相关性	线性回归系数	t 检验显著性	修正拟合度 R ²
锥形水泥杆				
杆长 15 m ⁺	0.243	4.505	0.383	-0.013
常数	-	91.696	0.036	
杆长 13 m ⁻	0.975	16.946	0.000	0.678
常数	-	2.320	0.000	
钢管杆				
不限杆型	0.369	6.644	0.246	0.070
常数	-	44.432	0.176	

可预测的 13 m⁻锥形水泥杆, 当前周期协议库存余量在 683 根, 供应可靠性概率估算为 86%。如果管理者计划将供应可靠性目标提升至 95% (对应置信水平 10%), 根据二元决策模型推算需求量, 差额量为 141 根。

6. 总结

线性回归分析在引入区间估计后, 形成了“数量 - 概率”的二元决策模型。利用这种改进模型, 电网物资供应链计划决策, 可以更加清晰供应可靠性。高压开关柜案例中, 检修小车的供应可靠性最低, 管理者有针对性地优先补充采购, 促使供应链资源分配与可靠性保持一致, 增强了供应链的整体韧性。架空线路案例中, 钢管杆和部分规格水泥杆没有通过假设检验。实际上, 这两种杆塔的选型设计, 受到地形环境的综合影响, 而因素又难以简单量化, 表明线性回归方法也有其局限性。即便如此, 分析结果仍然指出, 两种杆型的需求可预见性很不足, 供应链决策中需要采取缩短采购周期、倾斜备货力度等措施, 对冲因不可预见性而产生的风险。

参考文献

- [1] 魏华. 电网物力集约化管理模式下的协议库存物资计划预测管理[J]. 江西电力, 2015, 39(6): 82-84+94.
- [2] 邵帆. 电网企业物资需求预测技术研究及应用[D]: [硕士学位论文]. 北京: 华北电力大学(北京), 2017.
- [3] 丁靖, 林明晖, 陈凌. 历史数据挖掘下电网建设物资需求预测方法[J]. 信息技术, 2021(9): 144-148+154.
- [4] 毕子健, 王翎颖. 电网物资需求预测方法研究[J]. 华北电力技术, 2015(10): 26-30.
- [5] 丁红卫, 王文果, 万良, 等. 基于 BP 神经网络的电网物资需求预测研究[J]. 计算机技术与发展, 2019, 29(6): 138-142.
- [6] 王竹君, 朱颖琪, 孙界平. 基于矩阵分解的电力物资需求预测[J]. 四川大学学报(自然科学版), 2019, 56(4): 639-644.
- [7] 谢哲. 基于库存分类思想的电网工程物资需求预测研究[D]: [硕士学位论文]. 广州: 广东工业大学, 2022.
- [8] 赵欣, 丁靖, 吴建峰. 数据预测模型在物资需求计划申报中的应用[J]. 电子技术与软件工程, 2020(15): 173-174.
- [9] 睦楷, 王语涵, 王少勇, 等. 基于多元回归分析法的电网应急物资需求预测方法[J]. 电子技术与软件工程, 2016(23): 195-197.
- [10] 刘兆君. 正态系数线性回归的区间估计模型[J]. 统计与决策, 2018, 34(21): 37-40.
- [11] 李良哲. 正态总体置信区间平均长度的比较[D]: [硕士学位论文]. 武汉: 华中科技大学, 2023.

-
- [12] 杨建红, 张怀雄, 邹清明. 带 t -误差的线性回归模型参数的精确置信域[J]. 云南民族大学学报(自然科学版), 2004(1): 16-20.
 - [13] 白旭超. 两类重要模型的统计分析与可靠性评估研究[D]: [博士学位论文]. 西安: 西北工业大学, 2022.
 - [14] 郭庆光. 若干情形下的模型平均置信区间[D]: [硕士学位论文]. 青岛: 青岛大学, 2022.