

# 需求预测综合模型及其与库存决策的集成研究<sup>①</sup>

倪冬梅, 赵秋红, 李海滨

(北京航空航天大学经济管理学院, 北京 100191)

摘要: 准确的预测有助于企业做出有效的决策, 包括生产计划、定价和促销决策等, 以减少库存、提高客户满意度和企业竞争力。以快速消费品为研究对象, 分析其需求影响因素, 建立了时间序列分析与多元回归整合的需求预测综合模型; 将此预测模型引入到库存决策中, 构建了基于库存成本最小的需求预测与库存决策集成模型, 并借助变邻域搜索算法获得模型的参数值; 最后, 选用实际数据, 验证了所构建的需求预测综合模型、物流需求预测与库存决策集成模型及其求解方法的有效性。

关键词: 快速消费品; 需求预测; 时间序列分析; 多元回归模型; 库存决策

中图分类号: F253.1 文献标识码: A 文章编号: 1007-9807(2013)09-0044-09

## 0 引言

需求预测是物流管理的关键环节。明确需求的影响因素及其结果有利于提高预测的准确性, 从而减少库存, 使企业更有效地安排生产, 作出信息含量更高的定价和促销等决策, 提高客户满意度和企业竞争力。

采用适当的预测方法把握市场需求信息并将其作为决策依据已成为企业获取竞争优势的有效手段。Jain<sup>[1]</sup>对 1 236 家企业使用的预测模型进行调查研究发现, 61.33% 采用时间序列分析法, 22.65% 采用因果分析法, 13.92% 采用判断分析法, 2.10% 采用其它方法。该研究指出, 许多企业认识到促销、特殊事件、经济和价格等因素对需求的重要影响, 相继采用或考虑采用因果分析模型。

为了使需求预测更为有效, 学者们提出了很多方法。Gardner<sup>[2]</sup>阐述了指数平滑预测方法的研究现状, 强调该方法的经济预测特点以及鲁棒 (robust) 特征, 同时指出其存在的不足与局限性, 包括未考虑变量间的因果关系等。Thomassey 和 Happiette<sup>[3]</sup>基于神经网络并应用聚类 and 分类工具

提出了辅助决策系统——神经性聚类和分类系统, 解决产品缺乏历史数据的中期销售预测问题。Chern 等<sup>[4]</sup>提出新产品预测系统, 以解决新产品的销售预测, 该系统包括模型的选择, 求解及预测等模块, 结果表明该系统优于移动平均法且较少介入主观因素。

Bates 和 Granger<sup>[5]</sup>在 20 世纪 60 年代证明, 两种无偏的单项预测组合优于每个单项预测, 自此, 国内外有关组合预测的研究已经取得很大进展。周传世和刘永清<sup>[6]</sup>提出的变权重组合预测技术及周宏和廖雪珍<sup>[7]</sup>提出的 Logit 组合预测技术, 通过调整组合权重提高了预测精度。曾勇等<sup>[8]</sup>研究的多步贝叶斯信息更新情况下针对不同偏性特征的无偏组合预测模型, 扩展了贝叶斯组合预测模型的已有成果。Jeong 等<sup>[9]</sup>考虑供应链管理对预测的要求, 构建了基于 canonical 遗传算法 (CGA) 和 guided 遗传算法 (GGA) 来确定参数值的线性回归模型。邹昊飞等<sup>[10]</sup>建立的基于 GMDH (group method of data handling) 和遗传算法的两阶段神经网络组合预测模型, 提高了神经网络的

① 收稿日期: 2012-12-20; 修订日期: 2012-01-17.

基金项目: 国家自然科学基金资助项目(71071007; 71271013).

作者简介: 倪冬梅(1981—), 女, 黑龙江甘南县人, 博士. Email: ndmy@163.com

全局收敛能力、速度和预测精度。

一些学者基于供应链管理思想研究库存决策与预测模型的集成方法。Chandra 和 Grabis<sup>[11]</sup> 将自回归模型应用于供应链多阶段的需求预测,提出了基于库存管理的物料需求计划预测方法,以减小订货量的方差,即减轻牛鞭效应。Liang 和 Huang<sup>[12]</sup> 建立了基于多级代理商(multi-agent)的需求预测模型,研究结果表明,该模型能够有效降低总成本,并缓解订货量的波动程度。Su 和 Wong<sup>[13]</sup> 考虑制造商生产以及批发商补货过程中的固定成本,在应用指数平滑法来预测需求的基础之上,提出( $k, T_0$ )和两阶段蚁群算法相结合的方法确定弥补需求波动的超额订货量  $k$  和最优的补货周期  $T_0$ 。Ng 等<sup>[14]</sup> 将回归分析与时间序列组合起来,预测投标价格指数,结果表明,组合预测的预测效果优于两个模型单独预测的效果。

通过上述文献可以看出,采用综合预测模型进行需求预测,并将库存决策与需求预测进行整合,这是当今集成物流系统的发展方向。但现有的研究文献中,或者着重于综合预测模型的构建,或者将已有的需求预测方法与各种库存模型相嵌套,在构建综合预测模型并将其与库存决策相集成方面的研究较少。此外,在综合预测模型的基础上,现有的文献多考虑具有普适性的各预测模型的整合,未能够针对一类产品的特点,分析并量化其需求影响因素,从而构建有效和针对性强的需求预测综合模型。

针对目前研究的不足,本文以快速消费品(fast moving consumer goods, FMCG)为研究对象,分析其需求特征和影响因素,建立了基于时间序列分析和多元回归的综合预测模型(synthesized forecasting model based on the time series and the multi-regression methods, FMTR)。在综合预测模型的基础上,考虑一类典型库存决策问题,本文构建了基于库存成本最小的需求预测和库存决策集成模型(forecasting model integrated with the inventory decision, FMIID),并通过变邻域搜索(variable neighborhood search, VNS)算法求解模型的参数值。在理论分析的基础上,本文采用实际案例数据对所构建的综合预测模型以及库存决策与需求预测集成模型进行了分析评价。计算结果表明,综合预测模型的预测精度(平均误差、平均绝对误差

率和均方误差)均优于单个预测模型;库存决策与需求预测集成模型的预测精度小于综合预测模型,但其库存成本远低于传统的需求预测与库存决策各自独立的决策模型。将本文的预测结果与文献[11]和文献[14]进行比较,在预测精度方面,本文的综合预测模型优于文献[11]和文献[14]的预测模型,在成本方面,本文的集成模型比文献[11]和文献[14]的预测模型有较大成本节约。

## 1 FMCG 需求影响因素

消费品通常可以分为快速消费品(FMCG)和耐用消费品(Durable Goods, DG)两大类。快速消费品是目前市场上竞争比较激烈的产品,如白酒、香烟、饮料、饼干等。经营者依靠消费者高频次的使用与消耗、通过规模的市场流通量来获得利润。与耐用消费品相比,FMCG 具有以下特征:消耗周期短;消耗后需要及时补充;保质期短,具有质量保障的要求;属于视觉化、感受化产品,消费者在购买时很容易受卖场气氛及产品宣传的影响;品牌忠诚度不高,消费者很容易在同类产品中转换不同的品牌;单位价值一般较低,属于公众生活必需或消费普及类产品;需求对价格比较敏感,价格的变动尤其是同类产品价格的变动,将会对需求量产生影响,总之,其可替代性很强<sup>[15,16]</sup>。

对各种商品来讲,影响需求量的因素很多,消费者的购买行为也是难以把握的,这就加大了需求预测的难度,而 FMCG 的快速消费、多频次购买以及感受性强等特性更是加大了对需求进行预测的难度。与对其它产品的预测相同,可以从宏观和微观两个层面来考虑影响 FMCG 需求的因素:

1) 宏观因素 主要包括经济和社会两大类。经济性因素包括 GDP 和人均 GDP 指数等,如恩格尔系数的分析可很好地应用于食品、日用品等 FMCG 的长期预测中;社会性因素包括政治形势、宗教信仰、社会风俗等。

2) 微观因素 可概括为竞争、经营策略和产品自身特点 3 大类。竞争因素包括可替代产品的价格、促销、广告等因素;经营策略包括产品定价、

顾客服务等;产品自身因素包括产品生命周期、季节性因素和产品可得性等。

本文主要针对 FMCG 的短期需求预测.因此,认为在预测期间,宏观因素不发生变化. FMCG 短期需求预测的主要影响因素可概括为:产品生命周期、外部竞争者(主要是替代品价格等因素)、价格、促销因素、广告、季节趋势和销售计划。

## 2 综合预测模型

由于受多种因素影响,FMCG 的销量波动一般比较剧烈,前后期的差距较大.但同其它产品一样,其销量也存在惯性特征,即受到前几期,特别是最近一期需求的影响.考虑到 FMCG 的需求受到外界因素和自身历史销量的共同影响,本节建立了时间序列与多元回归整合的综合预测模型。

### 2.1 时间序列模型

时间序列的预测方法主要有移动平均法、指数平滑法、ARIMA(autoregressive integrated moving average model)法、趋势外推法等.通过对比分析,ARIMA 基于完善统计理论,对数据是否线性趋势无要求,具有较好的特性,对季节时间序列也有很好的拟合效果.虽然 ARIMA 较复杂,但随着计算机技术的发展,该方法越来越易于采用.故本文将采用 ARIMA 作时间序列预测<sup>[1]</sup>。

### 2.2 时间序列和多元回归综合预测模型

结合第 1 部分对需求影响因素的分析,选取适当的影响因子,建立包含时间序列的多元回归预测模型.选取的影响因子包括历史需求量、价格因素、促销因素、广告、竞争因素、所处的生命周期阶段、季节性趋势、节假日因素.具体来讲,所建立的时间序列和多元回归综合预测模型的步骤为,采用 ARIMA 预测模型,得到相关参数及第  $t$  期的预测值  $\hat{D}_{at}$ ,将  $\hat{D}_{at}$  与其它影响因素一起纳入回归模型,建立综合预测模型,如图 1 所示。

由于各影响因素数值差异大,采用单纯的线性回归不能很好地拟合需求的复杂变化.此外,指标值小的因子的解释能力也可能被缩小.为了弥补上述不足,可采用线性—对数模型(line-log model)<sup>[17]</sup>来获得更好的拟合效果。

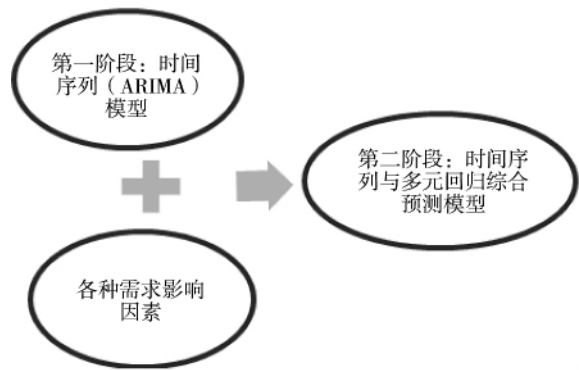


图 1 综合预测模型流程图

Fig. 1 The flowchart of two phases of the synthesized forecast model

基于综合预测模型的第 1 阶段 ARIMA 的预测值及考虑 FMCG 需求量的影响因素,可建立如下的多元回归综合预测模型框架

$$\hat{D}_t = \alpha_0 + \alpha_1 \ln \hat{D}_{at} + \sum_{k=1}^n \alpha_k x_{t,k} + \varepsilon_t \quad (1)$$

其中  $\hat{D}_t$  第  $t$  期回归模型预测值;  $\hat{D}_{at}$  ARIMA 预测值;  $x_{t,k}$  第  $t$  期第  $k$  个影响因素;  $\alpha_0, \alpha_1, \alpha_k$  待估参数;  $\varepsilon_t$  随机干扰因素,服从均值 0,方差  $\sigma^2$  的正态分布。

若第一阶段 ARIMA 所得模型为 AR 模型,即第  $t$  期的需求量与前  $p$  期需求量有相关关系,并将前  $p$  期的需求量作为多元回归的输入,则该模型变换为如下模型(2),可以得到更为精确的预测值

$$\hat{D}_t = \alpha_0 + \sum_{i=1}^p \alpha_{t-i} \ln D_{t-i} + \sum_{k=p+1}^n \alpha_k x_{t,k} + \varepsilon_t \quad (p < t) \quad (2)$$

其中  $\hat{D}_t$  第  $t$  期的需求量的预测值;  $D_{t-i}$  第  $t-i$  期的需求量;  $x_{t,k}$  第  $t$  期需求量的第  $k$  个影响因素;  $p$  自回归阶数,根据时间序列模型求得,  $p \leq t$ ;  $\alpha_{t-i}$  第  $t-i$  期销量的回归系数,待估计参数;  $\alpha_k$  第  $t$  期第  $k$  个影响因素的回归参数,待估计参数;  $\varepsilon_t$  随机干扰因素,服从均值 0,方差  $\sigma^2$  的正态分布。

具体回归分析中,可将模型(2)简化为

$$\hat{D}_t = \beta_0 + \beta_1 \ln D_{t-1} + \dots + \beta_p \ln D_{t-p} + \beta_{p+1} x_{t,p+1} + \dots + \beta_{v+n} x_{t,p+n} + \varepsilon_t \quad (3)$$

其中  $\beta_i$  为待估参数;其它同模型(2)。

根据历史数据,借助统计分析工具,如 Eviews,可分析所选影响因素的适当性,并可估计

出模型中的参数值.

### 3 需求预测与库存决策集成模型

不失一般性,假定每期期初,经销某种 FMCG 的某零售商根据自身库存及需求预测情况向分销中心订货,分销中心接到订单后可及时对零售商补货,分销中心的货源充足.由于 FMCG 的需求量较大,分销中心每期都要对零售商进行配送,相应的运输成本和订货成本为恒定的情形下,将不影响库存决策结果,因此,假设这两项成本为零.此外,由于产品具有被替代性强的特点,零售商处若缺货,将无法用下期补货来弥补.

在上述问题背景和假设条件下,库存决策模型的成本包括两项,一为供货量高于需求时的库存持有成本,一为供货量低于需求时的缺货成本.决策的目标是,确定每期的产品数量(为上期遗留库存量和本期订货量之和),以使库存成本和缺货成本总和的期望值达到最小.

#### 3.1 传统的需求预测与库存决策模型

传统上,库存决策和需求预测是单独进行的.结合本文所建立的需求综合预测模型,按照传统方法,需求预测和库存决策步骤如下.

首先,根据  $\min \sum_{t=1}^n (D_t - \hat{D}_t)^2$  原则,估计预测模型(1)的各项系数值.其中  $n$  为连续的历史数据的期数; $D_t$  为第  $t$  期的实际需求;  $\hat{D}_t$  为第  $t$  期的预测值.

其次,根据预测模型(1),求第  $t$  期的预测值  $\hat{D}_t$ . 本文假设需求服从正态分布,令第  $t$  期需求的均值和标准差分别为  $\mu_t = \hat{D}_t, \sigma_t = 1.25MAD_t$ . 其中,  $MAD$ (mean-absolute deviation) 是 Brown 在 1959 年提出的,可以通过指数平滑来估计,即

$$MAD_t = \alpha MAD_{t-1} + (1 - \alpha) | \hat{D}_{t-1} - D_{t-1} |$$

其中  $\alpha$  为平滑系数,一般取值 0.1 到 0.2 之间<sup>[18]</sup>.

最后,依据成本最小原则确定最优订货量.一般认为最优订货决策服从 order-up-to level(OUL, 依基准订货)策略<sup>[19]</sup>,则第  $t$  期的成本可以写成

$$\begin{aligned} \min C_t &= hI_t + wS_t \\ &= h [(\mu_t + Z\sigma_t) - D_t]^+ + \\ &\quad w [D_t - (\mu_t + Z\sigma_t)]^+ \end{aligned} \quad (4)$$

其中  $h$  为单位产品单位时间的库存成本;  $w$  为单位产品的缺货损失;  $S_t$  为第  $t$  期末的缺货量;  $I_t$  为第  $t$  期末的库存量,  $OUL_t = \mu_t + Z\sigma_t$  为第  $t$  期的库存水平,相应的订货量为  $Q_t = OUL_t - I_{t-1}$ ,如果  $a \geq 0$ , 则  $[a]^+ = a$ , 否则  $[a]^+ = 0$ , 则第  $t$  期的最优 OUL 以及相应的周期服务水平(cycle service level, CSL) 应满足

$$\begin{aligned} CSL^* &= \text{Prob}(D_t \leq OUL^* = \mu_t + Z^* \sigma_t) \\ &= \frac{w}{w + h} \end{aligned}$$

其中  $Z^* = F_z^{-1}\left(\frac{w}{w+h}\right)$ ,  $F_z^{-1}(\cdot)$  表示反标准正态分布函数值.

$n$  期的总成本(包括库存和缺货)即为

$$S_t = \sum_{t=1}^n \min C_t$$

#### 3.2 集成的需求预测与库存决策模型

不难看出,在单独决策下,库存决策是基于  $\min \sum_{t=1}^n (D_t - \hat{D}_t)^2$  所得到的需求预测结果,而该式的决策目标是需求预测误差最小,并没有考虑成本因素.当回归方程参数确定,预测值即确定,再根据预测值作库存决策,此时无法判定  $n$  期总成本是否最小.考虑到传统决策方法的不足,本文提出基于库存成本最小的回归方法来实现需求预测和库存控制的集成决策.

集成决策是指将需求预测和库存控制结合起来进行统一决策.本文提出的集成决策是依据  $n$  期总成本最小原则来估计回归参数并进行预测的,目标是直接找到每期初的最优目标库存水平  $OUL_t$ . 从上节的分析可以看出目标库存水平  $OUL_t = \hat{D}_t + Z\sigma_t$ , 如果假定需求量每期波动的方差是稳定的,那么可以认为  $OUL_t$  与各影响因素之间存在线性关系,可以用下式表示

$$\begin{aligned} OUL_t &= \beta_0 + \beta_1 \ln D_{t-1} + \dots + \beta_p \ln D_{t-p} + \\ &\quad \beta_{p+1} x_{t,p+1} + \dots + \beta_{p+n} x_{t,p+n} + \varepsilon_t \end{aligned} \quad (5)$$

第  $t$  期的库存成本可以写成

$$C_t = h(OUL_t - D_t)^+ + w(D_t - OUL_t)^+ \quad (6)$$

本文的优化目标是找到一组  $(\beta_0, \beta_1, \dots, \beta_{p+n})$  使  $n$  期的库存成本和缺货成本的总和最小, 即

$$\min h \sum_{i=1}^n (OUL_i - D_i)^+ + w \sum_{i=1}^n (D_i - OUL_i)^+ \quad (7)$$

这种集成决策过程如图 2 所示, 不需要对需求进行预测, 也不需要估计预测误差, 避免了因预测方法选择不当而造成的库存成本的上升. 在公式 (7) 中  $OUL_i$  为根据库存成本最小原则获得的回归预测值, 亦为最优的库存水平.  $h \sum (OUL_i - D_i)^+$  为库存成本,  $w \sum (D_i - OUL_i)^+$  为缺货成本, 则本期订货量为  $Q_i^* = OUL_i - I_{i-1}$ .

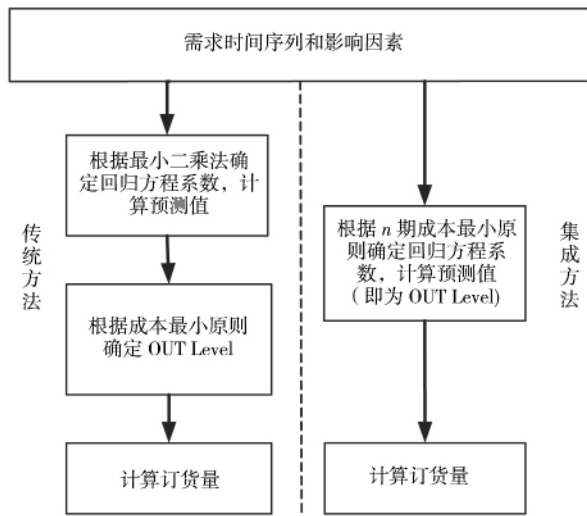


图 2 传统决策过程与集成决策过程对比图

Fig. 2 The contrast between the conventional and integrated decision processes

根据公式 (7), 当库存成本大于存货成本时, 为保持总成本最小, 目标函数会尽量减少库存水平, 使其小于需求; 反之, 则会尽量增加库存水平, 使其大于需求. 相比之下, 本文所提出的需求预测与库存决策集成模型更具灵活性. 根据快速消费品可替代性强的特点可知, 其缺货成本大于库存成本, 因此在总成本最小的目标约束下, 库存水平会尽量大于需求, 而传统的决策过程没有这样的调整过程, 因此传统决策过程下  $n$  期的总成本将不小于集成决策下的  $n$  期总成本.

### 3.3 集成模型的求解

集成模型参数较多, 求解难度大, 求解可采用

启发式算法. 本文借鉴 VNS 算法的思想求解集成模型的参数值. 算法可以在全局范围内搜索, 以获得问题的最优解或满意解<sup>[20]</sup>.

本文所采用的求解模型参数解的算法步骤如下.

1) 将综合预测模型得到的  $(\beta_0, \dots, \beta_n)$  作为初始解  $X_{cur}$ , 根据公式 (5) 计算初始解的目标函数值  $F_{cur}$ ;

2) 令当前最优解  $X_{opt} = X_{cur} = (\beta_0, \dots, \beta_n)$ ;

3) 令当前最优解的目标函数值  $F_{opt} = F_{cur}$ ;

重复以下步骤, 直至满足算法的停止规则  $N$ ,  $N$  为最大的迭代次数.

4) 设置  $k = 0$ ;

5) 直到  $k = k_{max}$ , 重复以下步骤:

① 若  $\beta_k > 0$ , 在区间  $[(1-a)\beta_k, (1+a)\beta_k]$  任取新值  $\beta'_k$ , 得到新的解  $X'_{cur}$ ; 若  $\beta_k < 0$ , 在区间  $[(1+a)\beta_k, (1-a)\beta_k]$  任取新值  $\beta'_k$ , 得到新的解  $X'_{cur}$ ;

② 计算  $X'_{cur}$  对应的目标函数值  $F'_{cur}$ . 如果  $F'_{cur} < F_{opt}$ , 则  $X_{opt} = X_{cur} = X'_{cur}$ ,  $F_{opt} = F'_{cur}$ ; 否则,  $k = k + 1$ .

## 4 案例分析

下面以某饮料公司在海南 44 周的饮料销售量 (单位: 箱) 为案例对象 (不失一般性, 本案例中, 将销售量近似看作需求量), 进行分析研究. 将 44 周的数据分为两大部分: 第 1 部分 (1 - 34 周) 作为模型训练区, 用于建立各个模型; 第 2 部分作为预测效果测试区 (35 - 44 周), 利用生成的模型对该区段的需求量进行预测. 根据产品特点, 设  $h = 0$  元,  $w = 5$  元.

### 4.1 综合预测模型验证

#### 4.1.1 时间序列的需求预测模型验证

需求原序列波动较大, 一阶差分后序列依然不平稳, 故对序列  $D_t$  作二阶差分得到序列  $D2_t$ . 当  $k = 1$  时, 样本自相关系数较大, 其余较小; 而且当  $k > 1$  时, 样本自相关系数的绝对值  $|\hat{\rho}_k| <$

$\frac{2}{\sqrt{N}}(1 + 2\hat{\rho}_1^2)^{1/2}$ , 故可以认为样本自相关系数在 1

步之后是截尾的,因此,可以用 MA(1) 模型对数据进行拟合.当  $k = 1$  时,样本偏相关系数较大,其余较小.样本偏相关系数的绝对值  $|\hat{\varphi}_{kk}| \leq \frac{2}{\sqrt{N}}$ ,故可以认为样本偏相关系数在 1 步之后是截尾的.因此,可以用 AR(1) 模型对数据进行拟合.结合样本自相关系数和偏相关系数的特点,根据 Box-Jenkins 建模思想可以尝试用 ARIMA(2, 2, 1) 模型进行拟合.结合 AIC 值、剩余平方和值及单位根的大小,根据简约原则可知,利用 AR(1) 对本案例中的需求时间序列进行拟合是比较恰当的.

利用 Eviews 软件对 AR(1) 模型参数进行检验,具体数值如表 1 所示.

从相伴概率可以看出,利用模型 AR(1) 对序列  $D_{2t}$  进行模拟比较恰当.对 AR(1) 进行适用性检验,残差序列的样本自相关和偏相关系数如表 2 所示.

表 1 AR(1) 模型参数检验

Table 1 The inspection of AR(1) model parameters

变量名	系数	标准误差	T 检验	概率
AR(1)	-0.668 903	0.133 831	-4.998 124	0.000 0
离差平方和	0.454 241	因变量均值	120.129 0	
调整离差平方和	0.454 241	因变量标准差	8 293.638	
回归标准差	6 126.966	Akaike 信息准则	20.310 51	
残差平方和	$1.13 \times 10^9$	Schwarz 准则	20.356 77	
最大似然估计	-313.813 0	D. W. 统计量	2.635 727	
特征根	-0.67			

表 2 残差序列的样本自相关系数图

Table 2 The autocorrelation and partial correlation of residuals series

自相关	偏相关	自相关系数	偏相关系数	Q 统计	概率
		1 -0.328	-0.328	3.663 7	
		2 -0.048	-0.174	3.745 6	0.053
		3 0.187	0.129	5.026 0	0.081
		4 -0.052	0.061	5.129 0	0.163
		5 -0.011	0.022	5.134 2	0.274
		6 0.029	0.000	5.168 6	0.396
		7 -0.021	-0.018	5.186 7	0.520
		8 0.077	0.080	5.448 4	0.605
		9 -0.090	-0.048	5.822 8	0.667
		10 0.036	0.001	5.887 4	0.751
		11 0.034	0.015	5.946 7	0.820
		12 -0.070	-0.032	6.207 0	0.859
		13 0.068	0.041	6.471 7	0.890
		14 0.001	0.022	6.471 8	0.927
		15 -0.006	0.031	6.473 7	0.953
		16 0.000	-0.015	6.473 7	0.971

从残差自相关系数可以看出,残差序列为纯

随机序列,所以 AR(1) 对序列  $D_{2t}$  是适应的.

综上所述,饮料需求量时间序列  $D_{2t}$  的模型为  $D_{2t} = -0.668 903D_{2t-1} + \mu_t$ ,进行差分逆运算可得

$$\hat{D}_t = 1.332 097D_{t-1} + 0.339 806D_{t-2} - 0.668 903D_{t-3}$$

#### 4.1.2 综合预测模型验证

从时间序列模型可以看出,第  $t$  期需求量  $D_t$  也受前 3 期需求量  $D_{t-1}$ 、 $D_{t-2}$  和  $D_{t-3}$  的影响,因此将前 3 期需求量  $D_{t-1}$ 、 $D_{t-2}$  和  $D_{t-3}$  作为 FMCG 的影响因素.由于所研究实例中的饮料具有稳定的广告策略,所研究的时段中并无新产品进入,且产品一直处于生命周期的成熟阶段,因此分析中广告因子、生命周期因子均为常数.此外,此需求数据来自于海南省饮料市场,季节性因素影响不大.所以,本案例中季节性因素、广告、生命周期因素均不计入实例的多元回归方程.结合实际中该饮料公司资料数据的可得性,经综合分析各因素对销量的影响,最终选取回归因子为价格、替代品价格变化率、节假日促销活动、 $D_{t-1}$ 、 $D_{t-2}$  和  $D_{t-3}$ .

综上所述,此例的综合预测模型为

$$\hat{D}_t = \beta_0 + \beta_1 \ln D_{t-1} + \beta_2 \ln D_{t-2} + \beta_3 \ln D_{t-3} + \beta_4 Prc_t + \beta_5 Cpt_t + \beta_6 Hda_t + \varepsilon_t$$

式中  $Prc_t$  为第  $t$  期的销售价格;  $Cpt_t$  为第  $t$  期的主要竞争信息,主要衡量关键竞争产品的价格变动信息,可表示为  $Cpt_t = (\text{替代品价格} - \text{价格}) / \text{价格}$ ;  $Hda_t$  为节假日指数,根据其对销量的影响力将其分等取值.根据公司访谈了解到,节假日促销对产品的影响力可由低到高赋予指标值为 1 和 2,非节假日其指标值为 0;为了描述节假日后的紧邻 1 期的需求低迷,引入指标值 0.8.当需求量大于 15 000 箱时,该饮料基本处于促销能力的第 3 级;当需求量大于 10 000 箱而小于 15 000 箱时,该饮料基本处于促销能力的第 2 级;当需求量大于 5 000 箱而小于 10 000 箱时,该饮料基本处于促销能力的第 1 级;当需求量小于 5 000 箱时,该饮料基本处于需求低迷期.

用 Eviews 求得各参数和回归残差,如表 3 和表 4.

由表 3 和表 4 可知,回归残差服从正态分布,说明该模型是适合的,所以案例中的多元回归模型为

表3 回归系数估计

Table 3 Regression coefficient estimation

变量名	系数	标准误差	T 检验	概率
C	-28 308.91	10 684.17	-2.649.614	0.012 4
DT1DS	2 368.480	669.591 0	3.537 204	0.001 3
DT2DS	1 409.505	644.642 5	2.186 490	0.036 2
DT3DS	326.747 9	586.432 9	0.557 179	0.581 3
HDA	2 365.512	458.720 0	5.156 768	0.000 0
PRC	-199.739 5	227.865 9	-0.876 566	0.387 3
CPT	35 542.83	12 154.33	2.924 293	0.006 3
离差平方和	0.795 011	因变量均值	7 523.821	
调整离差平方和	0.756 576	因变量标准差	4 491.196	
回归标准差	2 215.868	Akaike 信息准则	18.405 82	
残差平方和	$1.57 \times 10^8$	Schwarz 准则	18.704 41	
最大似然估计	-351.913 6	F 检验	20.684 35	
D.W. 统计量	1.430 189	F 检验概率	0.000 000	

表4 回归残差图

Table 4 The correlogram of regression residuals

自相关	偏相关	自相关系数	偏相关系数	Q 统计	概率
█	█	1 0.278	0.278	3.248 4	0.071
█	█	2 0.053	-0.026	3.370 8	0.185
█	█	3 0.037	0.031	3.430 1	0.330
█	█	4 0.029	0.013	3.469 1	0.483
█	█	5 -0.081	-0.101	3.774 0	0.582
█	█	6 -0.017	0.035	3.788 3	0.705
█	█	7 -0.018	-0.024	3.805 2	0.802
█	█	8 -0.041	-0.140	4.837 6	0.775
█	█	9 -0.184	-0.112	6.643 0	0.674
█	█	10 -0.313	-0.272	12.054	0.281
█	█	11 -0.298	-0.181	17.141	0.104
█	█	12 -0.175	-0.081	18.960	0.089
█	█	13 -0.091	-0.081	19.471	0.109
█	█	14 0.175	0.246	21.433	0.091
█	█	15 0.021	-0.124	21.463	0.123
█	█	16 -0.098	-0.149	22.129	0.139

表5 四种预测模型效果比较表

Table 5 The comparison of four forecast models

模型	平均误差 (ME)	平均绝对误差 (MAE)	平均绝对误差率 (%) (MAPE)
FMTR	$3.71 \times 10^{-12}$	1 563.40	0.27
RFMA	$-2.518 6 \times 10^{-12}$	2 028.24	0.38
文献[11]	-265.09	3 167.80	0.62
文献[14]	-93.58	2 419	0.42

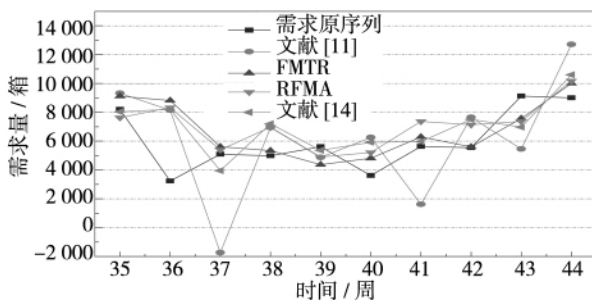


图3 需求量实际值和四个模型预测值曲线图

Fig. 3 The diagrams of the actual value and forecasting value of four models

$$\hat{D}_t = -28\,309 + 2\,368 \ln D_{t-1} + 1\,410 \ln D_{t-2} + 327 \ln D_{t-3} - 200 \text{Pr}_t + 35\,543 \text{Cpt}_t + 2\,366 \text{Hda}_t + \varepsilon_t$$

4.1.3 模型预测效果比较分析

下面对基于 ARIMA 预测值的综合预测模型 (将 ARIMA 预测值、价格、替代品价格变化率和节假日促销活动作为影响因素引入回归方程进行预测 (regression forecasting model based on ARIMA, RFMA)、基于原时间序列综合预测模型 (即本文所建立的综合预测模型 FMTR)、文献 [11] 和文献 [14] 4 个模型的预测效果从平均误差 (mean error, ME)、平均绝对误差 (mean absolute error, MAE) 和绝对误差百分率 (mean absolute percentage error, MAPE) 3 个角度进行比较 如表 5 所示。

从表5和图3可以看出,FMTR的ME、MAE和MAPE都小于文献[11]的,FMTR的ME大于RFMA的,但前者MAE和MAPE都小于后者的。可见,FMTR的预测精度比文献[11]高,FMTR的预测精度比RFMA高。

究其原因,综合预测模型不仅考虑了多期历史销量的影响因素,还考虑了需求的其它影响因素,FMTR避免了RFMA中ARIMA的预测误差。

4.2 需求预测与库存决策集成模型的算例

以35-44周的饮料需求量(单位:箱)及其预测值为案例对象,分别对传统库存决策模型和集成决策模型的决策结果进行分析研究。

在应用传统库存决策模型时,按照3.1节的步骤,根据本文所用资料信息,计算传统库存决策下的物流成本。

在应用集成决策模型时,采用 Matlab 软件编程,实现3.3节的VNS算法。设置  $k_{\max} = 6$  运行步数

$N = 5\ 000$  ,可得到  $F_{opt} = 5\ 278.6$  ,  $X_{opt} = (-26\ 842, 2\ 311, 1\ 494, 326, -204, 34\ 880, 2\ 355)$  ,故可得新的多元回归模型为

$$\hat{D}_t = -26\ 842 + 2\ 311 \ln D_{t-1} + 1\ 494 \ln D_{t-2} + 326 \ln D_{t-3} - 204Prct_t + 34\ 880Cpt_t + 2\ 355Hda_t + \varepsilon_t$$

对传统库存决策模型、需求预测与库存决策集成模型的计算结果进行比较分析 ,分别如图 4、表 6 和表 7 所示.

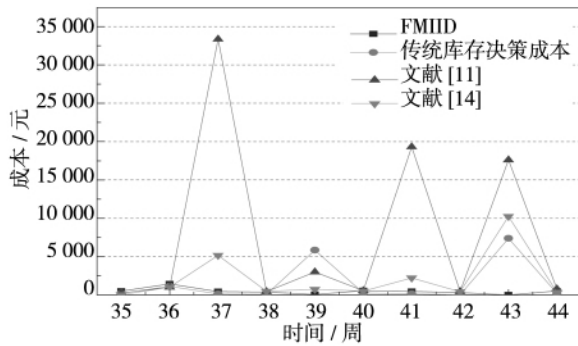


图 4 传统库存决策与集成库存决策成本曲线图

Fig. 4 The cost diagrams of the traditional and the integrated inventory decisions

表 6 传统与集成库存决策模型效果比较

Table 6 The components of the costs of traditional and the integrated forecasting models

模型名称	库存成本/元	缺货成本/元	总成本/元
FMTR 传统决策模型	2 055	13 998	16 053
FMIID	4 617	0	4 617
文献 [14]	4 937	16 150	21 087
文献 [11]	3 463	73 257	76 720

表 7 综合预测模型与库存集成模型预测效果比较

Table 7 The forecast outputs of the synthesized model and the integrated model

模型名称	平均误差 (ME)	平均绝对误差 (MAE)	平均绝对误差率 (MAPE)
FMTR	$3.71 \times 10^{-12}$	1 563.40	0.27
FMIID	417.36	1 641.77	0.29

从表 6 可以看出 ,需求预测与库存决策集成模型的转移成本为 4 617 元 ,大于传统模型下的 2 055 元 ,同时也大于文献 [11] 和 [14] 的传统决策成本. 通过前面分析可知 ,FMCG 所具有的特点决定了商家应该最大程度上避免缺货情况的发生. 需求预测与库存决策集成模型的缺货成本为 0 元 ,远小

于传统模型下的 13 998 元 ,从而前者的总成本也只有后者的 1/4.

从表 7 可以看出 ,集成模型预测结果的 3 个指标值都高于综合预测模型结果的相应指标值. 这是由于集成模型的预测是以成本最小而不是以预测误差最小为目标的 ,很显然 ,集成模型的预测精度低于传统模型的预测精度.

从以上分析结果以及两种回归预测方法的预测精度及其与库存成本的关系 ,不难得出结论: 在库存决策中 ,评价预测模型优劣的指标不应只是预测精度 ,还应包括基于此预测模型所做的库存决策的优劣——即库存成本的大小.

## 5 结束语

本文基于对 FMCG 需求影响因素的分析建立了基于 ARIMA 的时间序列和多元回归整合的综合需求预测模型 ,进一步地 ,从集成物流管理的角度出发 ,将该综合预测模型与库存决策相结合 ,构建了使物流成本最小的、需求预测与库存决策集成模型 ,并借助于 VNS 启发式算法 ,获得集成模型的参数值. 在案例分析部分 ,本文采用海南某饮料公司的 44 期的销售数据 ,以验证所构建的需求预测综合模型以及需求预测与库存决策集成模型的有效性. 计算结果显示 ,综合预测模型的预测精度高于单阶段模型的预测精度; 需求预测与库存决策集成模型的成本远低于非集成模型的成本 ,表明集成模型在减少缺货和节约成本方面具有明显的优势.

需求预测的内容复杂且广泛 ,需求预测与库存决策的集成是集成物流管理的重要研究方向 ,本文虽然在此方面作了一些工作 ,但还有许多问题值得继续深入研究. 在预测方法上 ,本文仅建立了 ARIMA 和多元回归综合预测模型 ,对其它预测模型的综合仍有待研究. 在需求预测和库存决策集成方面 ,进一步的研究应考虑不同背景下的库存模型以及多级物流系统的情形. 此外 ,本文所建立的需求预测与库存决策集成模型的决策目标是物流成本最小 ,当考虑其它因素 ,如服务水平要求时 ,如何建立和求解集成模型 ,也需要进一步探讨.



## 参考文献:

- [1] Jain C L. Benchmarking forecasting models[J]. *The Journal of Business Forecasting Methods and System*, 2002, 21(3): 18–20.
- [2] Gardner J E S. Exponential smoothing: The state of the art-Part II[J]. *International Journal of Forecasting*, 2006, 22(4): 637–666.
- [3] Thomassey S, Happiette M. A neural clustering and classification system for sales forecasting of new apparel items[J]. *Applied Soft Computing*, 2007, 7(4): 1177–1187.
- [4] Chern C C, Wu L L, Kung L C. Designing a decision-support system for new product sales forecasting[J]. *Expert Systems with Applications*, 2010, 37(2): 1654–1665.
- [5] Bates J N, Granger C W J. Combination of forecasts[J]. *Operations Research Quarterly*, 1969, (4): 451–468.
- [6] 周传世, 刘永清. 变权重组合预测模型的研究[J]. *预测*, 1995, 14(4): 47–48, 51.  
Zhou Chuanshi, Liu Yongqing. Study on the variable weight combination of forecasting model[J]. *Forecasting*, 1995, 14(4): 47–48, 51. (in Chinese)
- [7] 周宏, 廖雪珍. 市场需求 Logit 组合预测的研究[J]. *系统工程理论与实践*, 2003, 23(7): 63–69.  
Zhou Hong, Liao Xuezheng. Study on Logit combined forecasting of the market demand[J]. *Systems Engineering: Theory & Practice*, 2003, 23(7): 63–69. (in Chinese)
- [8] 曾勇, 唐小我, 郑维敏. 组合预测贝叶斯模型研究[J]. *管理科学学报*, 1999, 2(3): 14–21.  
Zeng Yong, Tang Xiaowo, Zheng Weimin. Bayesian models of combining forecasts[J]. *Journal of Management Sciences in China*, 1999, 2(3): 14–21. (in Chinese)
- [9] Jeong B, Jung H J, Spark N K. A computerized causal forecasting system using genetic algorithms in supply chain management[J]. *Journal of Systems and Software*, 2002, 60(3): 223–237.
- [10] 邹昊飞, 夏国平, 杨方廷. 基于两阶段优化算法的神经网络预测模型[J]. *管理科学学报*, 2006, 9(5): 28–35.  
Zou Haofei, Xia Guoping, Yang Fangting. Neural network forecasting model using multi-stage optimization approach based on GMDH and genetic algorithm[J]. *Journal of Management Sciences in China*, 2006, 9(5): 28–35. (in Chinese)
- [11] Chandra C, Grabis J. Application of multi-steps forecasting for restraining the bullwhip effect and improving inventory performance under autoregressive demand[J]. *European Journal of Operational Research*, 2005, 166(2): 337–350.
- [12] Liang W Y, Huang C C. Agent-based demand forecast in multi-echelon supply chain[J]. *Decision Support Systems*, 2006, 42(1): 390–407.
- [13] Su C T, Wong J T. Design of a replenishment system for a stochastic dynamic production/forecast lot-sizing problem under bullwhip effect[J]. *Expert Systems with Applications*, 2008, 34(1): 173–180.
- [14] Ng S T, Cheung S O, Skitmore M, et al. An integrated regression analysis and time series model for construction tender price index forecasting[J]. *Construction Management and Economics*, 2004, 22(5): 483–493.
- [15] Bottani E, Rizzi A. Economical assessment of the impact of RFID technology and EPC system on the fast-moving consumer goods supply chain[J]. *International Journal of Production Economics*, 2008, 112(2): 548–569.
- [16] Oke A, Long M. An analysis of the downstream logistics operations of a South African FMCG producer[J]. *International Journal of Production Economics*, 2007, 108(1/2): 176–182.
- [17] 李占风. 经济计量分析[M]. 北京: 中国财政经济出版社, 2005.  
Li Zhanfeng. *Analysis of Economic Instruments*[M]. Beijing: China's Financial Economy Press, 2005. (in Chinese)
- [18] Huh W, Janakiraman G, Muckstadt J A, et al. Asymptotic optimality of order-up-to policies in lost sales inventory systems[J]. *Management science*, 2009, 55(3): 404–420.
- [19] Brown R G. *Statistical Forecasting for Inventory Control*[M]. New York: McGraw-Hill, 1959.

(下转第74页)

## Classification of sentimental polarity for Chinese online reviews based on sentence level sentiment

WANG Hong-wei<sup>1</sup>, ZHENG Li-juan<sup>1</sup>, YIN Pei<sup>1</sup>, HE Shao-yi<sup>2</sup>

1. School of Economics and Management, Tongji University, Shanghai 200092, China;

2. College of Business Administration, California State University, San Marcos, USA

**Abstract:** With the boost of online reviews, sentiment polarity classification rises in response to the requirement of retrieving consumers' positive or negative opinions on certain products. The primary goal of this research is to improve the accuracy of sentiment polarity classification at the level of paragraphs for Chinese online reviews. With a view to the ways of expression and the grain of corpus, this paper presents a method to predict the sentiment polarity of Chinese online reviews in paragraphs based on sentence level sentiment analysis. Firstly, traditional classification methods are applied to predict the sentiment polarity of sentence. Then, three different algorithms i. e., the equal weight, correlation degree and assumption of sentiment condition, are employed to calculate the contribution that each sentence lying in the different positions of paragraph makes to the sentiment polarity of paragraph. Finally an experiment has been made based on hotel and mobile phone online reviews with lengths beyond two sentences. The result shows that the accuracy of sentiment polarity classification at the level of paragraph is remarkably increased by the method proposed in this paper, by taking correlation degree of expression and assumption of sentiment condition into consideration.

**Key words:** sentiment polarity; sentiment polarity of sentence; sentiment polarity of paragraph; Chinese online review; contribution to sentiment polarity

(上接第 52 页)

[20] Hansen P, Mladenović N, José A, et al. Variable neighborhood search: Methods and applications [J]. Operations Research, 2010, 175 (1): 367–407.

## Synthesized model for demand forecasting and its integration with inventory decision

NI Dong-mei, ZHAO Qiu-hong, LI Hai-bin

School of Economics and Management, Beihang University, Beijing 100191, China

**Abstract:** Accurate forecast is helpful for the enterprises to make decisions, including the production plan, the pricing and promotion decisions, etc., so as to reduce the inventory cost and improve the service qualities. In this paper, by analyzing the factors influencing the demand for fast moving consumer goods (FMCG), a synthesized forecasting model, which involves both the time series and the multi-regression methods, is established. The synthesized forecasting model is further integrated with the inventory decision, with the purpose of minimizing the overall logistics costs. To solve the integrated model where multiple parameters are involved, a variable neighborhood search (VNS) based algorithm is developed. To evaluate both the synthesized forecasting model and the integrated model of demand forecasting and inventory decision, computational studies are conducted based on some real data. The computational results show an outperformance of the synthesized forecasting model regarding forecasting accuracy, and an outperformance of the integrated model of demand forecasting and inventory decision when the logistics costs are minimized.

**Key words:** fast moving consumer goods (FMCG); demand forecast; time series analysis; multi regression model; inventory decision