

# 集成智能自适应货运量预测算法研究

于少伟

YU Shao-wei

山东英才学院 商学院, 济南 250104

Business School, Shandong Yingcai University, Jinan 250104, China

E-mail: yushaowei0505@126.com

**YU Shao-wei. Integrated intelligent and adaptive algorithm for freight transportation prediction. Computer Engineering and Applications, 2009, 45(32): 225-228.**

**Abstract:** A new uncertainty reasoning model based on cloud theory is firstly proposed after deeply analyzing the various freight transportation forecasting model, and then the relationship between freight transportation and related factors and the grey order of factors affecting freight transportation are analyzed and determined based on the theory of grey correlation. In addition, the freight transportation of Henan during 1997~2002 is predicted by a model based on one-dimension T-S-C-Y model and RBF neural network. The forecast model can not only integrate the fuzziness and randomness of a linguistic concept but also has the capacities of self-learning, self-organization, adaptation and strong fault-tolerant, and the output of the prediction model is the confidence interval which can well adapt the highly uncertain change of the freight transportation. The simulation results show that the outputs of T-S-C-Y model are in accord with the real ones, and the prediction model is more scientific and rational.

**Key words:** one-dimension T-S-C-Y model; cloud theory; freight transportation; Radial Basis Function(RBF); neural network

**摘要:** 在深入分析各种货运量预测模型的基础上, 首先提出了基于云理论的新的不确定性推理模型——T-S-C-Y模型; 然后基于灰色关联理论对货运量及其相关因素进行了灰关联分析, 并确定影响货运量诸多因素的灰关联序; 在此基础上结合一维T-S-C-Y模型和径向基函数神经网络建立了货运量预测模型对河南省1997~2002年的货运量进行预测。该预测模型既能很好地集成概念的模糊性和随机性, 同时又具有自学习、自组织、自适应和强的容错能力, 而且预测输出结果为置信区间值, 能够很好地适应货运量的高度不确定性。结果表明模型的预测值基本上与实际值吻合, 是一更加科学、合理的预测模型。

**关键词:** 一维T-S-C-Y模型; 云理论; 货运量; 径向基函数(RBF); 神经网络

**DOI:** 10.3778/j.issn.1002-8331.2009.32.070 **文章编号:** 1002-8331(2009)32-0225-04 **文献标识码:** A **中图分类号:** F224

随着我国社会经济的进一步发展, 货物运输需求将继续保持快速、稳定的增长趋势。而货物运输量的预测研究对于交通主管部门进行区域交通的整体规划, 加大交通设施体系的投资规模, 缓解区域交通运输的拥挤状况, 促进和保证该区域经济的持续、快速发展具有重要的现实意义<sup>[1]</sup>。

早期货运量的定量研究方法主要为指数平滑法和线性回归分析法等, 但由于这些方法仅适用于线性或弱非线性系统, 对于非线性系统的预测, 其精度难于满足预测要求<sup>[2]</sup>。

近年来随着灰色系统理论不断完善和发展, GM(1, 1)预测模型已越来越多地被应用于货运量的预测研究分析中<sup>[3-7]</sup>, 但运用GM(1, 1)模型预测时, 忽略了相关因素变化的影响。为此, 蔡家明<sup>[8]</sup>综合考虑了诸多因素对货运量的影响运用多维灰色模型对货运量变化进行了分析预测。在有关货运量的资料很难满足一般预测方法的要求的情况下, 灰色预测是一种比较简便的预测方法, 但该方法无法模拟人脑的思维结构。

人工神经网络是一种模拟人脑的思维结构, 具有自学习、自

组织、自适应能力和容错性强的定量预测法, 特别能有效地解决非线性预测问题。文献[8-9]采用神经网络来处理货运量的预测问题, 很好地模拟了人脑的思维, 但无法模拟人脑处理模糊和随机性问题的能力。

不同的预测方法提供不同的有用信息, 并且预测精度、侧重点往往也不同。因此可以将不同的预测方法采用一定的结构和参数进行适当的组合, 从而求得更精确的预测结果。文献[10-11]采用组合模型进行了预测, 并取得了很好的结果。但是过分注重精度, 忽略了结果的科学性和合理性以及与现实的吻合, 存在一定的不足。

上述模型都有各自的优势, 但它们存在一些共同的弱点: 这些方法给出一个或一系列的数据作为预测结果, 由于货运量的高度不确定性, 这些定量结果通常较脆弱, 且不易被非专业的用户所理解。

在深入分析了上述各种模型的基础上, 首先, 深入分析传统的基于云理论的不确定性推理模型并提出了基于云理论的

**基金项目:** 山东省高等学校科技计划项目(the Project of Shandong Province Higher Educational Science and Technology Program under Grant No.J09LG68)。

**作者简介:** 于少伟(1981-), 男, 讲师, 研究方向: 智能控制技术, 交通运输规划技术。

**收稿日期:** 2008-12-02 **修回日期:** 2009-07-20

新的不确定性推理模型——T-S-C-Y 模型,该模型可以有效地解决传统的基于云理论的不确定性推理模型在客观地反映现实中的复杂系统方面存在局限性的问题,可以逼近任意的非线性函数;然后,运用灰关联分析理论对影响货运量的各相关因素的主次关系进行深入、细致的分析研究;最后,结合一维 T-S-C-Y 模型和径向基函数神经网络建立了货运量预测模型并对货运量变化进行预测分析,从而能更全面、深入地了解诸多相关因素的变化及其对货运量变化的影响。该预测模型既能很好地模拟人脑处理模糊和随机性问题的能力并可以任意地逼近非线性函数,同时又具有自学习、自组织、自适应能力和强的容错能力,而且预测输出结果为置信区间值,能够兼顾预测精度以及预测值与现实的吻合,能够很好地适应货运量变化的高度不确定性,更具有科学性和合理性。

### 1 基于云理论的不确定性推理的局限

1995 年李德毅教授在传统模糊集理论和概率统计的基础上提出了既能表示和处理由模糊性引起的不确定性,又能表示和处理由随机性引起的不确定性的定性定量不确定性转换模型——云模型,并提出了不确定性推理的思想:如果一条定性规则的形式化描述为:If A then B; (A, B 为语言值对象)。则对照语言原子与云的对应关系,可以方便地运用带 X 条件的云对象和带 Y 条件的云对象来构造定性规则。对应一个固定的输入 U, 首先 X 条件云产生了一组随机的输出值  $\mu_i$ , 将输入变量的不确定性传递到输出变量空间, 而 Y 条件云在同一个  $\mu_i$  的控制下, 再次输出一组随机云滴(如图 1 所示)<sup>[12-13]</sup>。

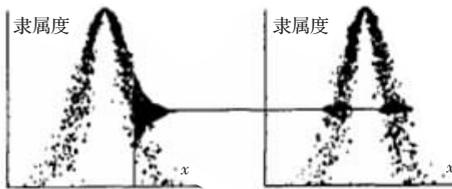


图 1 单规则发生机制

文献[14-22]中的基于云理论的不确定性推理的推理实现算法规定:“如果输入值激活的是规则前件的上升沿,则规则的后件也选择上升沿,反之亦然。”很明显该推理机制对现实中存在“如果输入值激活的是规则前件的上升沿,则规则的后件选择下降沿,反之亦然”的情况无能为力,尚不能任意地逼近非线性系统,存在一定的局限性,一定程度地影响了推理的精确性和科学性。

### 2 基于云理论的新的不确定性推理机制

鉴于上述局限性,在模糊 T-S 模型启发下,结合云理论的成功之处提出一种新的不确定性推理模型——T-S-C-Y 推理模型,模型把整个工作区域软化分成若干个云子区域,每一个云子区域对应的输出为输入值的线性表示,把输入值隶属于某个云子区域的隶属度作为该条规则的随机置信度,能够有效地弥补传统的基于云理论的不确定性推理模型在客观地反映现实中的复杂系统方面的不足,具有很强的函数逼近能力,能够大大降低推理过程的计算量,提高运算速度;而且可以与已有的线性系统理论很好地结合,可以与优化和自适应技术结合运用,非常适合于数学分析。

#### 2.1 一维 T-S-C-Y 推理模型

一维 T-S-C-Y 推理模型是一个多输入-单输出的系统,其

第 i 条推理规则的形式化描述为:

$$R_i: \text{if } x_1 \text{ is } A_{1i}, x_2 \text{ is } A_{2i}, \dots, x_m \text{ is } A_{mi}, \\ \text{then } y_i = P_{0i} + P_{1i}x_1 + P_{2i}x_2 + \dots + P_{mi}x_m$$

其中,  $x_j$  是第 j 个输入变量, m 为输入变量的数量,  $A_{ji}$  是一个云子集,其数字特征称为前件参数,采用基于黄金分割率的方法生成;  $y_i$  是第 i 条规则的输出,  $P_{ji}$  为后件参数,可以通过神经网络或者遗传算法等方法生成。

##### 2.1.1 单规则推理

一条一维两条件推理规则的形式化描述为: If  $x_1$  is  $A_1$  and  $x_2$  is  $A_2$  then  $y = p_0 + p_1x_1 + p_2x_2$ , 其中  $A_1, A_2$  为用云模型表示的语言值,称之为云子集。

图 2 为一维两条件单规则发生器示意图,图中  $CG_1, CG_2$  表示对应输入语言值  $A_1, A_2$  的 X 条件云,  $y = f(x_1, x_2)$  表示对应输入的线性输出。当某一特定的输入  $x_1$  和  $x_2$  刺激  $CG_1, CG_2$  时,  $CG_1, CG_2$  随机地产生一个二维隶属度向量, 然后进行软与运算得到  $\mu_k, \mu_k$  反映了对应该条定性规则的激活强度, 最后由规则后件线性输出。

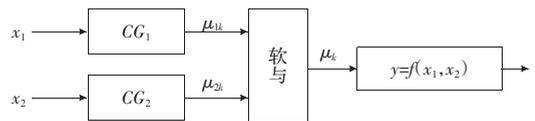


图 2 一维两条件单规则发生器

将前件云发生器和规则后件算法结合起来,得到如下推理算法:

- (1)  $En'_{1k} = G(En_1, He_1), En'_{2k} = G(En_2, He_2)$  分别生成以  $En_1$  为期望值、 $He_1$  为标准差的正态随机数  $En'_{1k}$  和以  $En_2$  为期望值、 $He_2$  为标准差的正态随机数  $En'_{2k}$ 。
- (2) 计算  $u_{1k} = \exp[-\frac{(x_1 - Ex_1)^2}{2En'^2_{1k}}]$  和  $u_{2k} = \exp[-\frac{(x_2 - Ex_2)^2}{2En'^2_{2k}}]$ 。
- (3) 将  $u_{1k}$  和  $u_{2k}$  软与得  $\mu_k$ , 作为该条规则的随机置信度。
- (4) 计算  $y = f(x_1, x_2)$ 。
- (5) 重复(1)至(3)步,直到得到足够的云滴。

##### 2.1.2 多规则推理

一维两条件多规则的形式化描述为: if  $x_1$  is  $A_{1i}, x_2$  is  $A_{2i}$  then  $y_i = P_{0i} + P_{1i}x_1 + P_{2i}x_2$ , 其中  $A_{1i}, A_{2i}$  为用云模型表示的语言值。

将前件云发生器和规则后件算法结合起来,得到一维两条件多规则推理算法:

- (1)  $En'_{i1k} = G(En_{i1}, He_{i1}), En'_{i2k} = G(En_{i2}, He_{i2})$  分别生成以  $En_{i1}$  为期望值、 $He_{i1}$  为标准差的正态随机数  $En'_{i1k}$  和以  $En_{i2}$  为期望值、 $He_{i2}$  为标准差的正态随机数  $En'_{i2k}$ 。
- (2) 计算  $u_{i1k} = \exp[-\frac{(x_1 - Ex_{i1})^2}{2En'^2_{i1k}}]$  和  $u_{i2k} = \exp[-\frac{(x_2 - Ex_{i2})^2}{2En'^2_{i2k}}]$ 。
- (3) 将  $u_{i1k}$  和  $u_{i2k}$  软与得  $u_{ik}$ , 作为该条规则的随机置信度。
- (4) 计算  $y_i = f_i(x_1, x_2)$ , 作为该规则的输出。
- (5) 通过公式  $y_k = \sum_{i=1}^n y_i * u_{ik} / \sum_{i=1}^n u_{ik}$  得到系统的随机输出。
- (6) 重复(1)至(4)步,直到输出值个数满足要求。
- (7) 然后由基于云 X 信息的逆向云新算法<sup>[23]</sup>得到整个推理系统的输出值  $y$ 。

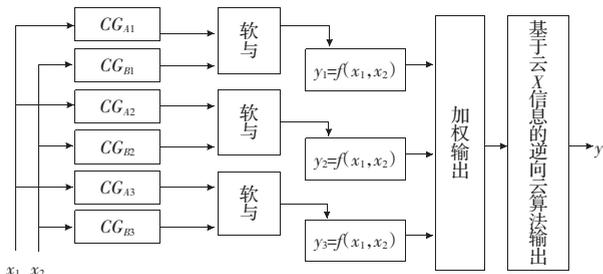


图3 一维两条件三规则推理模型

### 3 径向基函数神经网络

径向基函数(Radial Basis Function, RBF)网络<sup>[24]</sup>是在借鉴生物局部调节和交叠接受区域知识的基础上提出的一种采用局部接受域来执行函数映射的人工神经网络。

与BP网络相比具有如下优势:

(1)RBF网络隐层节点采用径向基函数,它仅在输入空间中一个很小的区域取非零值,但有不同的中心和宽度,对输入的反应有不同的兴奋程度,能在更高层次上实现对生物神经系统的模拟。

(2)对于层数相同的RBF网络,其表达能力往往强于BP网络。

(3)RBF网络的训练算法可以分为两段,各自都能快速化,无BP网络的缺点。

(4)RBF网络是一种典型的局部逼近网络,其学习速度快,泛化性能较好。

将RBF网络应用于复杂的预测会取得很好的效果。

## 4 基于一维T-S-C-Y模型和RBF网络的自适应预测模型

### 4.1 货运量驱动因子灰色关联分析

影响货运量的因子十分广泛且复杂,选取国内生产总值、总人口数、消费品零售总额、能源产量、铁路里程、公路里程6项指标作为综合货运量的影响因子,以它们1978~2001年的统计数据为基础进行货运量驱动因子灰色关联分析<sup>[25]</sup>(如表1所示)。

由表1可得出影响货运量的灰关联序为能源产量、公路里程、总人口数、铁路里程、零售品销售总额、生产总值,根据系统科学中的一条观察世界、分析事物的基本定理——哈肯伺服定理:任何复杂事物,不论它怎样复杂,对其起支配作用的因素(或变量)通常只有一个,或极少数的几个<sup>[26]</sup>。可以剔除影响较小的因子,而将所计算出的灰关联度较大的因子作为主要影响因素,来构建预测模型,以达到既简化预测模型又不影响预测结果的双重目的。

### 4.2 基于黄金分割率的云规则前件生成

基于黄金分割率的云生成方法的基本思想是<sup>[12]</sup>:给定的属性(即论域)看成语言变量,每个语言变量有几个语言值,语言值用云模型来表达,越接近论域的中心,云的熵和超熵越小,越远离论域的中心,云的熵和超熵越大,相邻云的熵和超熵的较小者是较大者的0.618倍。

在该研究中,采用基于黄金分割率的云生成方法在对货运量影响较大的3个变量各自的扩展论域上分别生成5朵隶属云,然后根据需要分别选择了其中的3朵隶属:  $A_1(7\ 067, 414.661, 1.6)$ ,  $A_2(8\ 311, 256.260, 1)$ ,  $A_3(9\ 555, 414.661, 1.6)$ ;  $B_1(4\ 434, 687.345, 1.6)$ ,  $B_2(6\ 496, 424.780, 1)$ ,  $B_3(8\ 558, 687.345, 1.6)$ ;  $C_1$

表1 货运量及与各相关因素的关联度

年份	生产总值/亿元	总人口/万人	消费品零售总额/亿元	能源产量/万吨标准煤	铁路里程/km	公路里程/km	货运量/万吨
1978	162.92	7 067	71.79	4 434	3 212	31 549	18 206
1979	190.09	7 189	80.44	4 536	3 216	36 155	17 535
1980	229.16	7 285	96.04	4 402	3 192	36 423	17 048
1981	249.69	7 397	106.86	4 760	3 460	36 478	16 403
1982	263.30	7 519	113.82	4 998	3 401	36 912	19 847
1983	327.95	7 632	123.05	5 456	3 305	37 196	21 579
1984	370.04	7 737	146.85	5 981	3 342	37 704	23 908
1985	451.74	7 847	180.59	6 909	3 248	38 840	35 642
1986	502.91	7 985	198.32	7 261	3 344	39 286	36 436
1987	609.60	8 148	225.28	7 361	3 409	39 713	39 539
1988	749.09	8 317	283.25	7 624	3 358	40 622	38 357
1989	850.71	8 491	310.85	8 031	3 546	41 170	38 245
1990	934.65	8 649	314.31	8 071	3 536	43 150	38 111
1991	1 045.73	8 763	368.92	7 999	3 384	44 199	39 923
1992	1 279.75	8 861	470.30	8 058	3 486	45 049	44 018
1993	1 660.18	8 946	577.96	8 037	3 456	46 487	47 347
1994	2 216.83	9 027	790.17	8 085	3 350	47 704	50 988
1995	2 988.37	9 100	957.76	8 454	3 382	49 707	53 582
1996	3 634.69	9 172	1 194.76	8 757	3 426	50 907	55 920
1997	4 041.09	9 243	1 427.53	8 558	3 428	55 016	56 113
1998	4 308.24	9 315	1 565.88	8 080	3 461	57 172	58 150
1999	4 517.94	9 387	1 691.20	6 947	3 354	60 330	59 218
2000	5 052.99	9 488	1 869.80	6 591	3 354	64 453	60 678
2001	5 533.01	9 555	2 071.93	7 238	3 319	69 041	65 191
R	0.735 1	0.939 4	0.773 5	0.960 5	0.931 9	0.950 5	

注:表中数据为河南省数据,来源于河南省2007年统计年鉴。

(31 549, 6 248.629, 1.6),  $C_2(50\ 295, 3\ 861.653, 1)$ ,  $C_3(69\ 041, 6\ 248.629, 1.6)$ 作为预测模型的规则前件。

### 4.3 后件网络的训练

为了克服ANFIS在以下方面的不足:(1)前件参数的自由更新能减弱模糊隶属函数的概念表达能力,破坏模糊假定规则的知识表达;(2)固定前件参数将会限制系统的逼近能力,文献[27]提出了一个新的学习模型:利用模糊规则把输入空间分成模糊区域,神经网络只在相应的模糊子区域里逼近。

受其启发,采用基于黄金分割率的云概念生成方法来划分输入空间,并利用径向基函数神经网络使该预测模型具备学习功能,增强其函数逼近能力。以Matlab神经网络工具箱为平台来实现各云子区域的网络生成和训练。

### 4.4 预测算法

结合基于黄金分割率生成的规则前件和由径向基函数神经网络逼近的规则后件得到如下的预测推理算法:

(1) $En_{i1k}=G(En_{i1}, He_{i1})$ ,  $En_{i2k}'=G(En_{i2}, He_{i2})$ 分别生成以 $En_{i1}$ 为期望值、 $He_{i1}$ 为标准差的正态随机数 $En_{i1k}'$ 和以 $En_{i2}$ 为期望值、 $He_{i2}$ 为标准差的正态随机数 $En_{i2k}'$ 。

(2)计算 $u_{i1k}=\exp[-\frac{(x-Ex_{i1})^2}{2En_{i1k}^2}]$ 和 $u_{i2k}=\exp[-\frac{(x-Ex_{i2})^2}{2En_{i2k}^2}]$ 。

(3)将 $u_{i1k}$ 和 $u_{i2k}$ 软与得到 $u_{ik}$ ,作为该条规则的随机置信度。

(4)计算 $y_i=sim(net_i)$ ,作为该规则的输出。

(5)通过公式 $y_k=\sum_{i=1}^n y_i^* u_{ik} / \sum_{i=1}^n u_{ik}$ 得到系统的随机输出。

(6)重复(1)至(4)步,直到输出值个数满足要求。

(7)然后由基于云 $X$ 信息的逆向云新算法得到整个推理系统的输出值 $y$ 。

(8)通过区间估计得到模型输出的样本总体均值的置信区间作为最终的输出结果。

#### 4.5 自适应参数调整

由于各影响因子随着时间的变化而变化,根据现有的数据采用基于黄金分割率的方法生成的预测模型的作用范围有限,而且输入与输出变量之间的非线性关系也随时间变化而变化,需要根据经济的发展情况不断地在线调整预测模型的参数来适应经济的发展,在此,根据正态分布的 $3\sigma$ 原则得到如下判断标准:如果指标值小于 $(Ex_{A3}+3He_{A3}) \cap (Ex_{B3}+3He_{B3}) \cap (Ex_{C3}+3He_{C3})$ ,则保持参数不变;如果指标值大于 $(Ex_{A3}+3He_{A3}) \cup (Ex_{B3}+3He_{B3}) \cup (Ex_{C3}+3He_{C3})$ ,则重新采用基于黄金分割率的云概念生成方法来划分输入空间,并重新利用径向基函数神经网络对各个子区域进行函数逼近,以实现预测模型参数的在线自适应调整。

### 5 预测仿真与结果分析

该研究采用 MATLAB 7.0<sup>[28]</sup>实现了一维 T-S-C-Y 模型和 RBF 神经网络的自适应货运量预测系统。图 4 显示了 1997 年至 2002 年连续 6 年的预测实验结果,图中包括实际数据曲线和自适应的预测曲线。

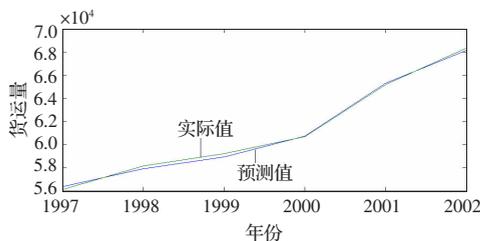


图4 连续6年的预测值与实际值

从图4可以看出预测值与实际值基本吻合,但是该预测模型的输出为大量服从正态分布的随机云滴,样本均值是一个随机值,存在一定的误差,且是一个单点值,尚不能很好地适应货运量的高度不确定性的变化。所以比较理想的是给出输出云滴总体分布的期望值的置信区间,讨论总体期望值的情况<sup>[29]</sup>。采用大子样对母体平均数区间估计的方法<sup>[30]</sup>,具体年份的母体均值的99%的置信区间如表2所示。

表2 预测结果的99%置信区间

年份	1997	1998	1999	2000	2001	2002
上限	56 655	58 222	59 286	60 907	65 610	68 415
下限	56 061	57 610	58 604	60 517	64 990	67 957

结合表1、表2和图5不难发现:实际的货运量值均落在预测结果的99%的置信区间上,更进一步验证了预测模型的科学性、准确性、合理性和客观性。

### 6 结论与研究展望

#### 6.1 结论

在模糊 T-S 模型启发下,结合云理论提出的 T-S-C-Y 模型,既能表示和处理由模糊性引起的不确定性,又能表示和处理由随机性引起的不确定性;各子区域对应的输出为输入值的线性表示,能够克服传统的基于云模型的推理机制在

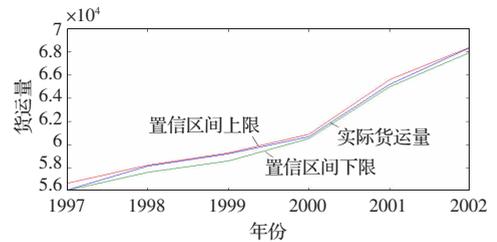


图5 连续六年的预测置信区间值与实际货运量值

方能逼近方面存在的不足,可以任意地逼近非线性函数,能够保证输出曲面的连续性,而且大大降低了推理过程的计算量,提高了运算速度;把输入值隶属于某个云子区域的隶属度作为该条规则的随机置信度,使输入值隶属于某个云子区域的隶属度的随机性在整个推理系统中传递,继承了云理论的成功之处。

基于一维 T-S-C-Y 模型和 RBF 神经网络的自适应货运量预测模型,既能表示和处理交通运输系统中由模糊性和随机性引起的不确定性,又能模拟人脑的思维结构,具有自学习、自组织、自适应能力和容错性强的特性,而且输出结果为置信区间值,能够很好地适应高度不确定的货运量变化,容易被非专业的用户理解。

#### 6.2 研究展望

基于一维 T-S-C-Y 模型和 RBF 神经网络的自适应货运量预测模型的规则前件采用基于黄金分割率的方法生成,其中的  $He$  和云滴数靠人为给定,存在一定的局限性,尝试建立一种既能很好地逼近规则后件又能同时调整  $He$  和云滴数的基于 T-S-C-Y 模型和神经网络的货运量预测模型。

基于一维 T-S-C-Y 模型和 RBF 神经网络的自适应货运量预测模型的输出结果为置信区间值,能够很好地适应高度不确定的货运量变化,容易被非专业的用户理解,但是单一预测模型的假设条件及适用范围总存在着这样或那样的局限性,且组合预测方法比单项预测的优越性早已经被 J.M Bates 和 C. W.J Granger 证明<sup>[31]</sup>,尝试研究基于区间输出的组合货运量预测模型。

#### 参考文献:

- [1] 陈唐民,李家本.汽车运输学[M].北京:人民交通出版社,1998.
- [2] 蔡家明.货运量多维灰色模型的预测[J].上海工程技术大学学报,2005,19(1):38-41.
- [3] 侯文英,李宣.包头市城区农产品物流园区规模预测[J].商场现代化,2008,526:82-83.
- [4] 李友俊,关欣.大庆市物流需求量的灰色预测[J].产业与科技论坛,2008,7(1):71-73.
- [5] 雷鸣涛.基于灰色系统理论的公路运输量预测[J].公路交通科技:应用技术版,2007(4).
- [6] 何国华.区域物流需求预测及灰色预测模型的应用[J].北京交通大学学报:社科版,2008,7(1):33-37.
- [7] 于德江.灰色系统建模方法的探讨[J].系统工程,1997(9):9-12.
- [8] 吴文兴.智能化物流预测模型的实现[J].现代计算机,2006,7:38-40.
- [9] 徐优丽.基于神经网络的物流需求预测—以浙江省长兴县为例[J].浙江树人大学学报,2008,8(1):56-58.
- [10] 程肖冰,张群.区域物流需求预测方法比较分析[J].工业工程与管理,2008,1:94-98.