

高阶直觉模糊时间序列预测模型

王亚男¹, 雷英杰¹, 雷阳², 范晓诗¹

(1. 空军工程大学防空反导学院, 陕西 西安 710051; 2. 武警工程大学电子技术系密码与信息安全武警部队重点实验室, 陕西 西安 710086)

摘要: 提出一种高阶直觉模糊时间序列预测模型。模型首先应用模糊聚类算法实现论域的非等分划分; 然后, 针对直觉模糊时间序列的数据特性, 提出一种更具客观性的直觉模糊集隶属度和非隶属度函数的确定方法; 最后, 利用直觉模糊多维取式推理建立高阶模型的预测规则, 进行预测。在 Alabama 大学入学人数和北京市日均气温 2 组数据集上分别与典型方法进行对比实验, 结果表明该模型有效提高了预测精度, 证明了模型的有效性和优越性。

关键词: 高阶; 直觉模糊时间序列; 预测模型; 直觉模糊推理

中图分类号: TP393.08

文献标识码: A

High order intuitionistic fuzzy time series forecasting model

WANG Ya-nan¹, LEI Ying-jie¹, LEI Yang², FAN Xiao-shi¹

(1. Air and Missile Defense College, Air Force Engineering University, Xi'an 710051, China;

2. Key Laboratory of CAPF for Cryptology and Information Security, Department of Electronics Technology, Engineering University of Armed Police Force, Xi'an 710086, China)

Abstract: A high order intuitionistic fuzzy time series forecasting model was built. In the new model, fuzzy clustering algorithm was used to get unequal intervals, and a more objective technique for ascertaining membership and non-membership functions of intuitionistic fuzzy set was proposed. On these bases, forecasting rules based on multi-dimension intuitionistic fuzzy modus ponens inference were established. At last, contrast experiments on the enrollments of the university of Alabama and the daily average temperature of Beijing were carried out, which show that the novel model has a clear advantage of improving the forecasting accuracy.

Key words: high order, intuitionistic fuzzy time series, forecasting model, intuitionistic fuzzy inference

1 引言

模糊时间序列 (FTS, fuzzy time series) 预测理论由 Song 和 Chissom 在 1993 年首次提出^[1]并应用于 Alabama 大学招生人数预测, 取得较好的预测效果^[2]。进而该理论在预测领域得到广泛研究与应用, 尤其在不精确数据、模糊数据或数据发展趋势不明确的问题中预测效果优良, 如粮食产量预测^[3]、股指交易量预测^[4]、环境污染浓度预测^[5]等。

对 FTS 预测模型的改良主要集中在 4 个方面。

1) 论域划分方法的改进。由研究初期的等分划分方法^[6]向预测精度更高的非等分划分方法^[7-10]改进。2) 模糊关系和预测规则的改进。预测规则的

建立以模糊关系为基础, 它们是 FTS 模型的核心部分。很多学者将神经网络方法^[11]、粒子群算法^[12]、矢量量化技术^[13]等引入 FTS 模型, 取得了比较精确的预测结果。3) 由一元一阶模型到多元和高阶模型的改进。高阶模型^[14,15]和多元模型^[16,17]的出现有效解决了一元一阶模型在实际应用上的局限性, 使 FTS 模型的适应性大大增强。4) 由非时变模型到时变模型的改进。时变模型中的模糊关系是一个随时间变化而变化的动态量, 这很大程度上提高了模型的适应性, 但计算量也随之增大^[18-21]。

在 FTS 模型中, 对历史数据模糊化时采用传统的 Zadeh 模糊理论, 即只用隶属度来度量数据的隶属性质, 这既不够全面也不够客观, 从根本上限制

收稿日期: 2015-07-23; 修回日期: 2015-10-19

基金项目: 国家自然科学基金资助项目 (No.61309022)

Foundation Item: The National Natural Science Foundation of China (No.61309022)

间序列。

3 基于直觉模糊推理的高阶 IFTS 预测模型

传统 IFTS 预测模型可以概括为以下 4 个步骤。

- 1) 确定论域大小，并对论域进行划分。
- 2) 建立直觉模糊集，将历史数据直觉模糊化。
- 3) 建立直觉模糊逻辑关系，确定预测规则，求得预测值。
- 4) 预测结果去模糊化输出。

本文的高阶 IFTS 预测模型将步骤 4) 的去直觉模糊化过程融合到步骤 3) 的预测计算中，将模型缩减为 3 个主要步骤。本节以通用数据集 Alabama 大学招生人数为例，介绍所建模型。

3.1 基于模糊聚类的非等分论域划分

定义问题的讨论范围，即论域 $U = [x_{\min} - e_1, x_{\max} + e_2]$ ，其中， x_{\min} 和 x_{\max} 分别表示历史数据的最小值和最大值； e_1 和 e_2 是 2 个合适的正数，为了讨论和计算上的方便使 x_{\min} 和 x_{\max} 分别向下、向上取合适的整数。在本例中，历史数据为 1971 年~1991 年共 21 年招生数，待预测数据为 1992 年招生数，其中 $x_{\min} = 13\ 055$ ， $x_{\max} = 19\ 337$ ，令 $e_1 = 55$ ， $e_2 = 663$ ，则论域 $U = [13\ 000, 20\ 000]$ 。

在论域划分上，FTS 模型的理论和方法在 IFTS 模型上同样适用，Chen 和 Huang 等^[6, 7]的研究已表明，在 FTS 模型中使用非等分划分方法会产生比等分方法更好的预测结果。近年来，很多学者采用遗传算法、粒子群算法和 FCM 等算法来实现非等分的论域划分，这类算法的每个子区间具有明显实际意义，更符合人们的理解习惯。但是，这类算法通常在大数据条件下才能充分发挥其优良性能，而这又与 FTS 和 IFTS 模型不需要大量历史数据的优势相背离；此外，IFTS 预测方法在实际中也多用于数据量不大的经济、环境等问题。因此，本文采取了较上述优化算法更简捷、更具实时性的基于最大生成树的直接模糊聚类算法^[27]。

设 $X = \{x_1, x_2, \dots, x_n\}$ 是待分类对象的全体，其中， $x_i = (x_{i1}, x_{i2}, \dots, x_{im})$ ， $i = 1, 2, \dots, n$ ，即 x_i 有 m 维特征。所谓最大树法就是以被分类对象为顶点，相似矩阵 R 的元素 r_{ij} 为边的权重而得到的一棵最大树，然后取定阈值 $l \in [0, 1]$ ，去掉权重小于 l 的边，得到一个不连通的图，其中，各连通分支的顶点便构成了一个在 l 水平上的分类。对于不同的 l 可以

得到不同的分类，从而得到一组动态分类，因此算法的最后还需确定一个最佳阈值 l 从而对应一个最优分类，确定最佳阈值的方法有 F 统计量方法等。文献[27]对算法具体步骤已有详细介绍，本文不再赘述。

通过聚类得到的 r 类数据可记为

$$\{x_1^1, x_2^1, \dots, x_{n_1}^1\}, \{x_1^2, x_2^2, \dots, x_{n_2}^2\}, \dots, \{x_1^{i-1}, x_2^{i-1}, \dots, x_{n_{i-1}}^{i-1}\}, \\ \{x_1^i, x_2^i, \dots, x_{n_i}^i\}, \{x_1^{i+1}, x_2^{i+1}, \dots, x_{n_{i+1}}^{i+1}\}, \dots, \{x_1^r, x_2^r, \dots, x_{n_r}^r\}$$

其中， $x_{n_{i-1}}^{i-1} < x_1^i$ ， $x_1^i < x_2^i < \dots < x_{n_i}^i$ ， $x_{n_i}^i < x_1^{i+1}$ ， n_i ($i = 1, 2, \dots, r$) 表示第 i 类数据所含对象个数。

记

$$d_i = \begin{cases} x_{\min} - e_1, & i = 0 \\ \frac{x_{n_i}^i + x_1^{i+1}}{2}, & i = 1, 2, \dots, r-1 \\ x_{\max} + e_2, & i = r \end{cases} \quad (6)$$

则论域 U 被划分成的 r 个非等长区间可表示为 $u_1 = [d_0, d_1], u_2 = [d_1, d_2], \dots, u_r = [d_{r-1}, d_r]$ 。

本例中得到历史数据的最优分类数为 9，进而将论域 U 划分为 9 个区间（为便于计算将所有数据均缩小 1 000 倍）。

$$u_1 = [13.000, 13.309], u_2 = [13.309, 14.282], \\ u_3 = [14.282, 14.921], u_4 = [14.921, 16.186], \\ u_5 = [16.186, 16.598], u_6 = [16.598, 17.535], \\ u_7 = [17.535, 18.560], u_8 = [18.560, 19.149], \\ u_9 = [19.149, 20.000].$$

3.2 历史数据直觉模糊化

对应于论域 U 的 9 个区间，定义 9 个代表语言变量的直觉模糊集

$$A_i = \{ \langle x, m_{A_i}(x), g_{A_i}(x) \rangle \mid x \in X \} \quad (7)$$

其中， $i = 1, 2, \dots, 9$ 。它们的现实意义可以理解为：“很少”、“少”、“较少”、“微少”、“正常”、“微多”、“较多”、“多”、“很多”。

确定 A_i 的隶属度和非隶属度函数是该步骤的关键所在。由于直觉指数的影响，确定直觉模糊集的隶属度和非隶属度函数的方法呈现极大的复杂性。已有方法如模糊统计法、三分法、二元对比排序法等都将直觉指数固定为一个常数^[28]，有失客观性，因此，本节针对 IFTS 模型实际应用情况和划分区间的特性，给出一种新的方法。

首先，通过客观分析，给出以下 2 条规则。

1) 当 x 位于区间中点即 $x = \frac{d_{i-1} + d_i}{2}$ 时, $m_{A_i}(x) = 1, g_{A_i}(x) = 0$ 。

2) 当 x 位于区间边界即 $x = d_i$ 时, 直觉指数取最大值, 隶属度和非隶属度相等, 根据具体情况令 $p_{A_i}(d_i) = a, (0 < a < 1)$, 则 $m_{A_i}(d_i) = g_{A_i}(d_i) = \frac{1-a}{2}$ 。

然后, 从上述规则出发, 确定隶属度函数采用典型 Gaussian 函数

$$m_{A_i}(x) = \exp\left(-\frac{(x - c_{mi})^2}{2s_{mi}^2}\right) \quad (8)$$

非隶属度函数由 Gaussian 函数变形得到

$$g_{A_i}(x) = 1 - \exp\left(-\frac{(x - c_{gi})^2}{2s_{gi}^2}\right) \quad (9)$$

则直觉指数函数为

$$p_{A_i}(x) = 1 - m_{A_i}(x) - g_{A_i}(x) \quad (10)$$

其中, $i = 1, 2, \dots, r$, c_{mi} 、 s_{mi} 和 c_{gi} 、 s_{gi} 为函数的参数, 是确定隶属度和非隶属函数的关键, 依然由上述规则计算得到

$$c_{mi} = c_{gi} = \frac{d_{i-1} + d_i}{2} \quad (11)$$

$$s_{mi}^2 = \frac{-(d_i - c_{mi})^2}{2 \ln \frac{1-a}{2}} \quad (12)$$

$$s_{gi}^2 = \frac{-(d_i - c_{gi})^2}{2 \ln \left(1 + \frac{1-a}{2}\right)} \quad (13)$$

定义 8 A 为有限论域 X 上的直觉模糊集, 如果 A 满足:

- 1) $0 \leq m_A(x) \leq 1, 0 \leq g_A(x) \leq 1$;
- 2) $0 \leq p_A \leq 1, 0 \leq m_A(x) + g_A(x) \leq 1$;
- 3) $m_A(x) + g_A(x) + p_A(x) = 1$ 。

则称直觉模糊集 A 为正规直觉模糊集。

定理 1 本节提出的隶属度函数、非隶属度函数和直觉指数函数的确定方法是规范的, 即 A_i 是正规直觉模糊集。

证明 1) 由 Gaussian 函数的定义, 有 $0 \leq \exp\left(-\frac{(x-c)^2}{2s^2}\right) \leq 1$, 即 $0 \leq 1 - \exp\left(-\frac{(x-c)^2}{2s^2}\right) \leq 1$ 。故得 $0 \leq m_{A_i}(x) \leq 1, 0 \leq g_{A_i}(x) \leq 1$ 。

2) 由 $0 < a < 1$ 且 $\ln \frac{1-a}{2} < \ln \left(1 + \frac{1-a}{2}\right)$, 有 $s_{mi}^2 < s_{gi}^2$, 又因为 $c_{mi} = c_{gi}$, 代入 Gaussian 函数中, 可得 $\exp\left(-\frac{(x - c_{mi})^2}{2s_{mi}^2}\right) > \exp\left(-\frac{(x - c_{gi})^2}{2s_{gi}^2}\right) > 0$ 。又由 $0 \leq \exp\left(-\frac{(x - c_{mi})^2}{2s_{mi}^2}\right) \leq 1, 0 \leq \exp\left(-\frac{(x - c_{gi})^2}{2s_{gi}^2}\right) \leq 1$, 有 $1 - \exp\left(-\frac{(x - c_{mi})^2}{2s_{mi}^2}\right) < \exp\left(-\frac{(x - c_{gi})^2}{2s_{gi}^2}\right) \leq 1$ 。故得 $1 - \exp\left(-\frac{(x - c_{mi})^2}{2s_{mi}^2}\right) < \exp\left(-\frac{(x - c_{gi})^2}{2s_{gi}^2}\right) > 0$, 即 $0 \leq \exp\left(-\frac{(x - c_{mi})^2}{2s_{mi}^2}\right) + 1 - \exp\left(-\frac{(x - c_{gi})^2}{2s_{gi}^2}\right) \leq 1$ 。因此, $0 \leq m_{A_i}(x) + g_{A_i}(x) \leq 1$ 。

进而, 得 $0 \leq p_{A_i}(x) = 1 - m_{A_i}(x) - g_{A_i}(x) \leq 1$ 。

3) 由直觉指数函数的确定方法, 明显可得 $m_{A_i}(x) + g_{A_i}(x) + p_{A_i}(x) = 1$ 。

证毕。

本例中, 令 $a = 0.4$, 计算得到各参数的值如表 1 所示。

表 1 隶属度和非隶属度函数参数值

直觉模糊集	$c_{\mu i}$	$s_{\mu i}$	$c_{\gamma i}$	$s_{\gamma i}$
A_1	13.154 5	0.009 9	13.154 5	0.033 5
A_2	13.795 5	0.098 3	13.795 5	0.331 8
A_3	14.601 5	0.042 4	14.601 5	0.143 1
A_4	15.553 5	0.166 2	15.553 5	0.560 8
A_5	16.392 0	0.017 6	16.392 0	0.059 5
A_6	17.066 5	0.091 2	17.066 5	0.307 7
A_7	18.047 5	0.109 1	18.047 5	0.368 2
A_8	18.854 5	0.036 0	18.854 5	0.121 6
A_9	19.574 5	0.075 2	19.574 5	0.253 8

各直觉模糊集的隶属度函数、非隶属度函数和直觉指数如图 1~图 3 所示。

最后, 将历史数据直觉模糊化。计算历史数据对每个直觉模糊集的隶属度、非隶属度和直觉指数的值, 其中, 最大隶属度对应的直觉模糊集即该历史数据所属的直觉模糊集。本例中 21 个历史数据各自对应的直觉模糊集如表 2 所示。

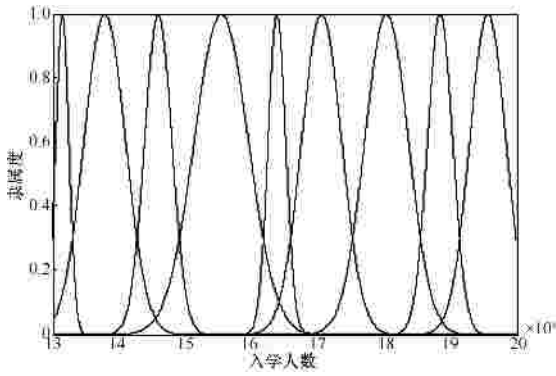


图 1 各直觉模糊集的隶属度函数

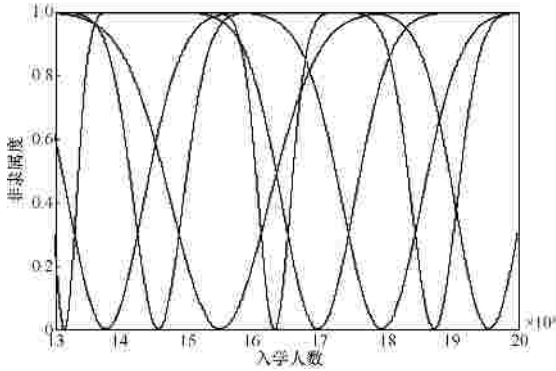


图 2 各直觉模糊集的非隶属度函数

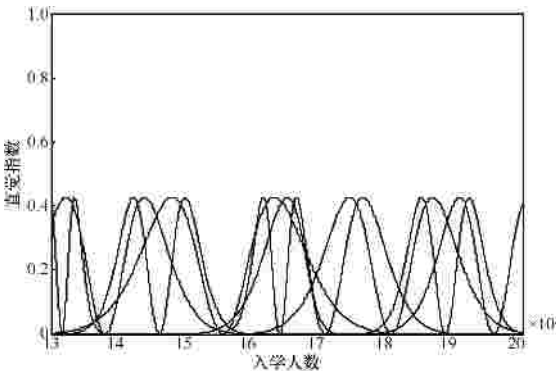


图 3 各直觉模糊集的直觉指数函数

表 2 历史数据对应的直觉模糊集

年份	直觉模糊集	年份	直觉模糊集
1971	A_1	1982	A_4
1972	A_2	1983	A_4
1973	A_2	1984	A_4
1974	A_3	1985	A_4
1975	A_4	1986	A_4
1976	A_4	1987	A_6
1977	A_4	1988	A_7
1978	A_4	1989	A_8
1979	A_6	1990	A_9
1980	A_6	1991	A_9
1981	A_5		

3.3 直觉模糊逻辑关系和预测规则

本文采用启发式直觉模糊多维取式推理进行预测，首先介绍多维取式推理的基本概念。

3.3.1 直觉模糊多维取式推理

设 $A_i (i=1, 2, \dots, n)$ 和 $A_i^* (i=1, 2, \dots, n)$ 是 U 上的直觉模糊集， B 和 B^* 是 V 上的直觉模糊集，基于直觉模糊关系的广义多维取式推理规则^[29]陈述的是：给定两直觉模糊命题“ x 是 $A_1 \times A_2 \times \dots \times A_n$ ”和“如果 x 是 $A_1 \times A_2 \times \dots \times A_n$ ，则 y 是 B ”，可推出一个新的直觉模糊命题“ y 是 B^* ”，且 $A_1^* \times A_2^* \times \dots \times A_n^*$ 与 $A_1 \times A_2 \times \dots \times A_n$ 近似， B^* 与 B 近似。推理模型为

规则：IF x is $A_1 \times A_2 \times \dots \times A_n$ THEN y is B

输入： x is $A_1^* \times A_2^* \times \dots \times A_n^*$

输出： y is B^*

令 $A = A_1 \times A_2 \times \dots \times A_n$ ， $A^* = A_1^* \times A_2^* \times \dots \times A_n^*$ ，则模型中的表述规则可简记为直积空间 $U \times V$ 上的直觉模糊关系 $A \rightarrow B$ ，其关系矩阵 $R = R(A; B)$ 是一个双矩阵。在直觉模糊推理中，关系矩阵的计算运用不同的算子，可得到不同的 m_R 与 g_R ，所以计算结果的数值不唯一，而推理输出的结果基本是一致的。其中，Mamdani 蕴涵算子性能较好且便于计算，因此这里选其进行推理计算^[28]，即

$$R = A \times B = \int_{U \times V} \frac{\langle m_R(x, y), g_R(x, y) \rangle}{(x, y)} \quad (14)$$

其中，

$$m_R(x, y) = m_A(x) \wedge m_B(y) \quad (15)$$

$$g_R(x, y) = g_A(x) \vee g_B(y) \quad (16)$$

$$m_A(x) = \bigwedge_{i=1}^n m_{A_i}(x) = m_{A_1}(x) \wedge m_{A_2}(x) \wedge \dots \wedge m_{A_n}(x) \quad (17)$$

$$g_A(x) = \bigvee_{i=1}^n g_{A_i}(x) = l_{A_1}(x) \vee g_{A_2}(x) \vee \dots \vee g_{A_n}(x) \quad (18)$$

则推理输出为

$$B^* = A^* \circ R \quad (19)$$

其中，直觉模糊合成运算“ \circ ”采用“取大—取小”运算，即

$$m_{B^*}(y) = \bigvee_{x \in U} (m_{A^*}(x) \wedge m_R(x, y)) \quad (20)$$

$$g_{B^*}(y) = \bigwedge_{x \in U} (g_{A^*}(x) \vee g_R(x, y)) \quad (21)$$

将式(24)、式(27)和式(28)的前件和后件分别以对应的直觉模糊集 F_j 替代,可得变形后的直觉模糊逻辑关系为

$$F_{19}, F_{20}, F_{21} \rightarrow F_i \quad (30)$$

$$F_{17}, F_{18}, F_{19} \rightarrow F_{20} \quad (31)$$

$$F_{18}, F_{19}, F_{20} \rightarrow F_{21} \quad (32)$$

分别将式(31)和式(32)作为 3.3.1 节中直觉模糊多维推理的推理规则,式(30)的前件作为推理输入,推理输出分别记为 F_i^1 和 F_i^2 ,则推理模型分别为

规则: IF x is $F_{17} \times F_{18} \times F_{19}$ THEN y is F_{20}

输入: x is $F_{19} \times F_{20} \times F_{21}$

输出: y is F_i^1

和

规则: IF x is $F_{18} \times F_{19} \times F_{20}$ THEN y is F_{21}

输入: x is $F_{19} \times F_{20} \times F_{21}$

输出: y is F_i^2

按式(14)~式(23)进行推理合成运算,可以得到 F_i^1 中拥有最大隶属度的为 A_7 、 A_8 和 A_9 这 3 个连续的元素, F_i^2 中拥有最大隶属度的为 A_8 。

Step3 推理结果去直觉模糊化。

直觉模糊集的解模糊算法通常有重心法、加权平均法和最大真值法等,由于重心法具有

比较平滑的输出控制,及对应于输入信号的微小变化输出也会发生一定的变化,且这种变化明显比较平滑,因此本节采用重心法对推理结果解模糊化^[28]。重心法的具体计算方法是取隶属度函数和非隶属度函数合成的真值函数曲线与横坐标围成的平面图像面积的重心为解模糊的输出值,即

$$C_0 = \frac{\int_U c \left(m_F(c) + \frac{\pi_F(c)}{2} \right) dc}{\int_U \left(m_F(c) + \frac{\pi_F(c)}{2} \right) dc} \quad (33)$$

其中, U 为输出论域, F 为定义在论域 U 上的直觉模糊子集。

计算可得 F_i^1 去直觉模糊化的结果为 $f_1 = 18.97$, F_i^2 去直觉模糊化的结果为 $f_2 = 18.7675$ 。

Step4 推理结果加权合成输出。

当只存在一个推理结果,即在 Step1 中只找到一个相似关系时,那么该推理结果的权重 $w = 1$,将该推理结果作为最终预测结果输出。当存在多个推理结果时,则需要分别计算各推理结果 f_i 的权重 w_i ,将它们的加权和 $f = \sum w_i f_i$ 作为最终结果输出。下面以 F_i^1 、 F_i^2 为例介绍具体计算方法。

1) 采用基于标准化欧式距离的相似度计算方法,如式(34)所示,分别计算式(30)与式(31)的前件相似度 s_1 ,式(30)与式(32)的前件相似度 s_2 。

$$s = 1 - d(A, B) = 1 - \sqrt{\frac{1}{2n} \sum_{i=1}^n \left[(m_A(x_i) - m_B(x_i))^2 + (g_A(x_i) - g_B(x_i))^2 + (p_A(x_i) - p_B(x_i))^2 \right]} \quad (34)$$

$$s_1 = \frac{3 - d(F_{19}, F_{17}) - d(F_{20}, F_{18}) - d(F_{21}, F_{19})}{3} = 0.6161 \quad (35)$$

$$f = w_1 f_1 + w_2 f_2 = 18.8793 \quad (39)$$

$$s_2 = \frac{3 - d(F_{19}, F_{18}) - d(F_{20}, F_{19}) - d(F_{21}, F_{20})}{3} = 0.7623 \quad (36)$$

至此,就得到了 1992 年的招生人数的预测结果为 18 880 人。

4 模型应用

4.1 Alabama 大学数据集实验

按照本文所建模型继续对 Alabama 大学招生数据集的其他各年数据进行预测,以检测其可行性,所得结果如表 4 所示。

根据均方误差 (RMSE, root mean square error) 和平均预测误差 (AFE, average forecasting error) 2 项指标对本文模型的预测性能进行检验,指标的计算如式(40)和式(41)所示,计算结果如

2) 计算权重 w_1 、 w_2

$$w_1 = \frac{s_1}{s_1 + s_2} = 0.447 \quad (37)$$

$$w_2 = \frac{s_2}{s_1 + s_2} = 0.553 \quad (38)$$

3) 推理结果加权合成

表 5 所示。

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\text{真实值}_i - \text{预测值}_i)^2}{n}} \quad (40)$$

$$AFE = \frac{\sum_{i=1}^n \left| \frac{\text{真实值}_i - \text{预测值}_i}{\text{真实值}_i} \right|}{n} \times 100\% \quad (41)$$

表 4 本文模型对 Alabama 大学入学人数的预测结果

年份	实际招生数	预测值
1971	13 055	—
1972	13 563	—
1973	13 867	—
1974	14 696	14 233
1975	15 460	15 400
1976	15 311	15 420
1977	15 603	15 513
1978	15 861	15 605
1979	16 807	16 780
1980	16 919	17 010
1981	16 388	16 513
1982	15 433	15 605
1983	15 497	15 600
1984	15 145	15 221
1985	15 163	15 300
1986	15 984	15 853
1987	16 859	16 807
1988	18 150	17 545
1989	18 970	18 238
1990	19 328	18 780
1991	19 337	18 975
1992	18 876	18 880

将文献[1, 16, 23, 25]中的模型在该数据集上的预测性能一同列入表 5 中进行比较。其中,文献[1]为传统 FTS 模型;文献[16]为高阶 FTS 预测模型,文中作者选择的最优预测结果所对应的“200-4-second order”模型将论域划分为 35 个区间,这在实际操作中会带来极大的工作量,是不现实的,因此本文在保证计算合理的基础上选则其性能最优的“300-3-first order”模型进行比较;文献[23]和文献[25]为一阶 IFTS 预测

模型,为现有 IFTS 预测模型中并没有高阶模型,所以只能与一阶模型进行比较。

表 5 各模型对 Alabama 大学入学人数的预测性能

指标	RMSE	AFE/%
文献[1]	677.1	3.35
文献[16]	433.6	2.56
文献[23]	418.9	2.07
文献[25]	399.3	1.99
本文	303.1	1.27

由表 5 可知,本文所建高阶 IFTS 模型能够有效预测通用数据集中的数据,预测效果较现有模型有了较大提升,表明模型具有很强的可行性和有效性。

4.2 日均气温数据集实验

北京市日均气温数据集是由中国气象信息中心提供的以天为单位对北京市气温进行的统计,取每日 4 次定时(02:00、08:00、14:00、20:00)观测值的平均值为该日气温平均值,本文只选取从 2014 年 6 月 1 日~2014 年 7 月 1 日的 31 个数据作为实验数据集,如表 6 所示。

表 6 2014.6.1~2014.7.1 北京市日均气温(单位:0.1)

日期	气温	日期	气温
6.1	247	6.17	232
6.2	220	6.18	240
6.3	252	6.19	240
6.4	257	6.20	221
6.5	280	6.21	220
6.6	217	6.22	220
6.7	231	6.23	250
6.8	233	6.24	273
6.9	233	6.25	244
6.10	217	6.26	262
6.11	237	6.27	292
6.12	278	6.28	299
6.13	260	6.29	288
6.14	259	6.30	298
6.15	262	7.1	288
6.16	259		

在该数据集上应用文献[1]、文献[16]、文献[23]和文献[25]中的模型及本文模型进行预测,预测值及实际值如图 4 所示。

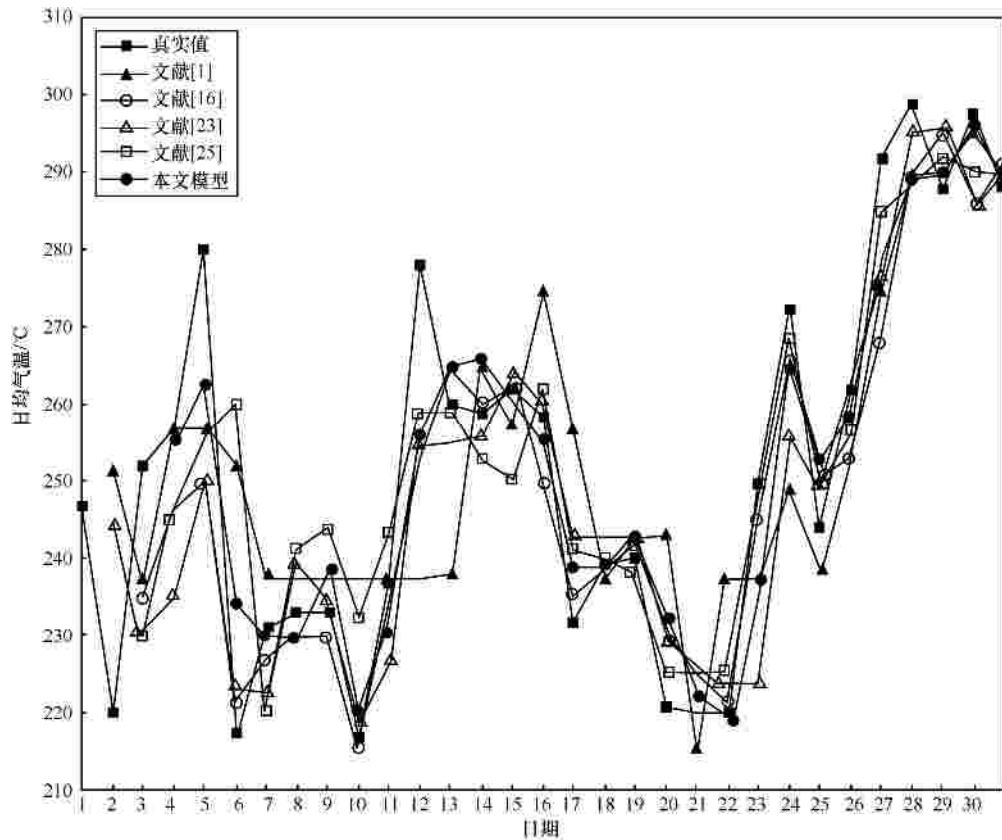


图 4 各模型对日均气温数据集的预测结果

各模型预测性能对比如表 7 所示。

表 7 各模型对日均气温数据集的预测性能

指标	RMSE	AFE/%
文献[1]	16.98	5.3
文献[16]	10.62	2.9
文献[23]	12.94	3.92
文献[25]	12.96	3.91
本文	9.39	2.89

5 结束语

本文针对模糊时间序列预测方法的不足，利用直觉模糊集在处理不确定数据集上的优势，建立了一个新的高阶直觉模糊时间序列预测模型。采用基于最大支撑树的模糊聚类算法实现论域的非等分划分，使论域划分这一基础步骤更具实时性和简捷性；针对划分数据的实际特性，给出一种更客观的直觉模糊集隶属度和非隶属度的确定方法；采用直觉模糊多维取式推理建立序列数据的预测规则，更有效地预测不确定数据的模糊变化特征。在 2 类数据集上与经典算法的对比实验表明模型具有较好

的预测性能。如何建立多元高阶直觉模糊时间序列，进一步提高预测性能，将是下一步研究的重点。

参考文献：

- [1] SONG Q, CHISSOM B S. Forecasting enrollments with fuzzy time series-Part I[J]. Fuzzy Sets and Systems, 1993, 54(1): 1-9.
- [2] SONG Q, CHISSOM B S. Fuzzy time series and its models[J]. Fuzzy Sets and Systems, 1993, 54(1): 269-277.
- [3] SINGH S R. A robust method of forecasting based on fuzzy time series[J]. Applied Mathematics and Computation, 2007, 188(1): 472-484.
- [4] QIU W, LIU X, WANG L. Forecasting shanghai composite index based on fuzzy time series and improved c-fuzzy decision trees[J]. Expert Systems with Applications, 2012, 39(9): 7680-7689.
- [5] DOMANSKA D, WOJTYLAK M. Application of fuzzy time series models for forecasting pollution concentrations[J]. Expert Systems with Applications, 2012, 39(9): 7673-7679.
- [6] CHEN S M. Forecasting enrollments based on fuzzy time series[J]. Fuzzy Sets and Systems, 1996, 81(3): 311-319.
- [7] HUANG K, YU T H K. Ratio-based lengths of intervals to improve fuzzy time series forecasting[J]. IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, 2006, 36(2): 328-340.
- [8] LU W, CHEN X, PEDRYCZ W, et al. Using interval information granules to improve forecasting in fuzzy time series[J]. International Journal of Approximate Reasoning, 2015, 57: 1-18.
- [9] LIU J W, CHEN T L, CHENG C H, et al. Adaptive-expectation based

- multi-attribute FTS model for forecasting TAIEX[J]. Computers and Mathematics with Applications, 2010, 59(2): 795-802.
- [10] CHENG C H, CHEN T L, WEI L Y. A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting[J]. Information Science, 2010, 180(9): 1610-1629.
- [11] KHASHEI M, HEJAZI S R, BIJARI M. A new hybrid artificial neural networks and fuzzy regression model for time series forecasting[J]. Fuzzy Sets and Systems, 2008, 159(7): 769-786.
- [12] PARK J I, LEE D J, SONG C K, et al. TAIEX and KOSPI 200 forecasting based on two-factor high-order fuzzy time series and particle swarm optimization[J]. Expert Systems with Applications, 2010, 37(2): 959-967.
- [13] LI S T, KUO S C, CHENG Y C, et al. Deterministic vector long-term forecasting for fuzzy time series[J]. Fuzzy Sets and Systems, 2010, 161(13): 1852-1870.
- [14] TSAI C C, WU S J. A study for second-order modeling of fuzzy time series[C]//1999 IEEE International Fuzzy Systems Conference Proceedings. Seoul, Korea, c1999: 22-25.
- [15] AVAZBEIGI M, DOULABI S H H, KARIMI B. Choosing the appropriate order in fuzzy time series: a new n -factor fuzzy time series for prediction of the auto industry production[J]. Expert Systems with Applications, 2010, 37(8): 5630-5639.
- [16] ALADAG C H, BASARAN M A, EGRIOGLU E, et al. Forecasting in high order fuzzy times series by using neural networks to define fuzzy relations[J]. Expert Systems with Applications, 2009, 36(3): 4228-4231.
- [17] CHEN S M, CHEN S W. Fuzzy forecasting based on two-factor second-order fuzzy-trend logical relationship groups and the probabilities of trends of fuzzy logical relationships[J]. IEEE Transactions on Cybernetics, 2015, 45(3): 405-417.
- [18] SONG Q, CHISSOM B S. Forecasting enrollments with fuzzy time series-part II[J]. Fuzzy Sets and System, 1994, 62(1): 1-8.
- [19] HWANG J R, CHEN S M, LEE C H. Handling forecasting problems using fuzzy time series[J]. Fuzzy Sets and Systems, 1998, 100(1): 217-228.
- [20] CHEN S M, HWANG J R. Temperature prediction using fuzzy time series[J]. IEEE Transactions on Systems, Man, and Cybernetics-part B: Cybernetics, 2000, 30(2): 263-275.
- [21] LIU H T, WEI N C, YANG C G. Improved time-variant fuzzy time series forecast[J]. Fuzzy Optimization Decision Making, 2009, 8(1): 45-65.
- [22] CASTILLO O, ALANIS A, GARCIA M, et al. An intuitionistic fuzzy system for time series analysis in plant monitoring and diagnosis[J]. Applied Soft Computing, 2007, 7(4): 1227-1233.
- [23] JOSHI P, KUMAR S. Intuitionistic fuzzy sets based method for fuzzy time series forecasting[J]. Cybernetics and Systems: An International Journal, 2012, 43(1): 34-47.
- [24] 郑寇全, 雷英杰, 王睿, 等. 直觉模糊时间序列建模及应用[J]. 控制与决策, 2013, 28(10): 1525-1530.
- ZHENG K Q, LEI Y J, WANG R, et al. Modeling and application of IFTS[J]. Control and Decision, 2013, 28(10): 1525-1530.
- [25] 郑寇全, 雷英杰, 王睿, 等. 参数自适应的长期 IFTS 预测算法[J]. 系统工程与电子技术, 2014, 36(1): 99-104.
- ZHENG K Q, LEI Y J, WANG R, et al. Method of long-term IFTS forecasting based on parameter adaptation[J]. Systems Engineering and Electronics, 2014, 36(1): 99-104.
- [26] ATANASSOV K. Intuitionistic fuzzy sets[J]. Fuzzy Sets and Systems, 1986, 20(1): 87-96.
- [27] 梁保松, 曹殿立. 模糊数学及其应用[M]. 北京: 科学出版社, 2007: 75-77.
- LIANG B S, CAO D L. Fuzzy mathematics and applications[M]. Beijing: Science Press, 2007: 75-77.
- [28] 雷阳. 基于直觉模糊核匹配追踪的目标识别方法研究[D]. 西安: 空军工程大学, 2012.
- LEI Y. Research on target recognition techniques based on intuitionistic fuzzy sets and kernel matching pursuit[D]. Xi'an: Air Force Engineering University, 2012.
- [29] 雷英杰, 赵杰, 路艳丽, 等. 直觉模糊集理论及应用(上册)[M]. 北京: 科学出版社, 2014: 145-155.
- LEI Y J, ZHAO J, LU Y L, et al. Theories and applications of intuitionistic fuzzy set[M]. Beijing: Science Press, 2014: 145-155.

作者简介:



王亚男(1988-),女,山东青岛人,空军工程大学博士生,主要研究方向为网络信息安全。



雷英杰(1956-),男,陕西渭南人,空军工程大学教授、博士生导师,主要研究方向为网络信息安全、智能信息处理。

雷阳(1984-),女,陕西西安人,博士,武警工程大学副教授,主要研究方向为网络信息安全。

范晓诗(1988-),男,陕西西安人,空军工程大学博士生,主要研究方向为网络信息安全。