# 震害快速评估中基于GM(1,1)模型的人口预测

周中红，何少林，陈文凯
（中国地震局兰州地震研究所，甘肃 兰州 730000）

**摘 要**：地震灾区人口数据是破坏性地震发生时震害快速评估中一项重要的关键基础数据。本文利用GM(1,1)模型建立了甘肃省县一级人口数量的增长模型，实现了静态人口基础数据在地震发生时的动态计算，使其更接近于人口发展的实际情况，为震害评估提供更准确的人口基础数据。

**关键词**：震害快速评估；人口预测；GM(1,1)模型；甘肃
**中图分类号**：P315.9；C921 **文献标识码**：A **文章编号**：1000－0844(2009)01－0071－04

# Application of GM(1,1) Model for Population Forecasting in Fast Assessment of Earthquake Disaster

ZHOU Zhou-hong, HE Shao-lin, CHEN Wen-kai
(*Lanzhou Institute of Seismology, CEA, Lanzhou* 730000, *China*)

**Abstract**: Population data in stricken-region is a fundamental and significant aspect in fast assessment of destructive earthquake. In this paper, a county level population growth model for Gansu province is set up by means of GM(1,1), and dynamic computing is operated based on static population data to better fit the actual situation, in order to provide more accurate population foundation data to fast earthquake disaster assessment.

**Key words**: **Fast assessment of earthquake disaster; Population forecasting; GM(1,1) model; Gansu province**

## 0 引言

人员伤亡的估计是破坏性地震发生时震害快速评估中最重要的任务之一，直接决定地震紧急救援的规模和行为，而人口分布与增长数据的不确定性是影响地震人员伤亡快速评估的重要因素[1]。在当前的地震人员伤亡快速评估系统中，人口库的数据是静态的，并没有考虑人口增长变化的规律。而地震的发生是随时的，使得评估模型中使用的陈旧人口数据与实际情况有较大的偏差，导致人员伤亡评估结果的不准确。虽然人口数据举足轻重，但由于数据涉及其它行业、部门，及时更新的费用高、周期长、更新难度很大。针对此问题，本文从建立人口预测模型的角度对人口基础数据进行改善，使其更接近于人口发展的实际情况，为震害评估提供更准确的人口基础数据。

人口预测有许多模型可选，较为常用的有一元线性回归模型、马尔萨斯人口模型、logistic模型、GM(1,1)灰色模型等，每种模型都有其适用范围的不同、繁简程度的差异。如一元线性回归模型要求样本容量较大；马尔萨斯人口模型主要针对有些地区人口增长短期内没有明显的约束，人口发展前一段时期较慢，越往后发展速度越快的情况下选用；Logistic模型一般适用于人口增长率开始下降的情况，并且该模型主要用于远期预测等[2-4]。而GM(1,1)灰色预测方法的主要特点是建立预测模型所需要的原始数据不多，特别适合数据难以求得的场合，使得该预测方法简便并具有较高的准确性[5-10]；同时可用编程语言实现适用于不同的人口指标(如人口总数、出生人口数、死亡人口数、自然增长人口

收稿日期：2008-05-14
基金项目：中国地震局"十五"重点项目"甘肃省地震应急指挥系统"；中国地震局兰州地震研究所论著编号：LC2009002
作者简介：周中红(1975－)，女(汉族)，甘肃临夏人，在读硕士，主要从事震害评估方法研究工作。

数等)、不同范围、不同起止年间的原始数据序列,可对未来任意年间的人口数据进行预测,并且预测时间越近,误差越小[9]。由于人口发展是一个比较典型的灰色系统,加之笔者想要实现的是在有限的样本容量下较为短期的预测结果,因此选用了GM(1,1)灰色模型建立甘肃省87个县区一级行政单元的人口增长预测模型。可根据此模型实现静态人口基础数据在地震发生时的动态计算、改善人口数据库更新缓慢的问题,提高人口数据库的精度。

# 1 GM(1,1)模型的建模

灰色理论是我国学者邓聚龙教授于20世纪80年代前期提出的用于控制和预测的新理论、新技术[11]。可通过少量的、不完全的信息寻找其内在规律,建立灰色模型(Grey Model,简称GM),对系统的未来状态做出科学的定量预测。GM(1,1)模型是基于灰色系统理论的最常用的预测模型[11-12]。

## 1.1 GM(1,1)的建模步骤

(1) 对原始数据序列 $\boldsymbol{X}^{(0)}=[x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n)]$ 做1-AGO(1次累加),得到 $\boldsymbol{X}^{(1)}=[x^{(1)}(1),x^{(1)}(2),\cdots,x^{(1)}(n)]$,其中

$$x^{(1)}(t)=\sum_{k=1}^{t}x^{(0)}(k) \qquad (1)$$

(2) 对于 $\boldsymbol{X}(1)$ 建立灰色预测模型GM(1,1),即建立白化方程

$$\frac{\mathrm{d}x^{(1)}}{\mathrm{d}t}+ax^{(1)}=b \qquad (2)$$

其中:灰色参数 $a$,$b$ 分别为发展系数、内生控制灰数。该方程的解为

$$\hat{x}^{(1)}(t+1)=\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-at}+\frac{b}{a} \qquad (3)$$

经还原,有

$$\hat{x}^{(0)}(t+1)=\hat{x}^{(1)}(t+1)-\hat{x}^{(1)}(t) \qquad (4)$$

## 1.2 误差检验

灰色系统理论采用三种方法检验、判断模型的精度。

(1) 残差大小检验:对模型值和实际值的误差进行逐点检验,其中残差

$$\varepsilon(k)=\boldsymbol{X}^{(0)}(k)-\hat{x}^{(0)}(k) \qquad (5)$$

相对误差

$$\Delta_k=\frac{|\varepsilon(k)|}{\boldsymbol{x}^{(0)}(k)}\times 100 \qquad (6)$$

对于给定的相对误差 $\alpha$,当 $\bar{\Delta}<\alpha$ 且 $\Delta_k<\alpha$ 成立时,称模型为残差合格模型。

(2)关联度检验:通过考察模拟值曲线与建模序列曲线的相似程度进行检验,则关联度

$$r_0=\frac{1}{n}\sum_{k=1}^{n}\varepsilon_0(k) \qquad (7)$$

其中 $\varepsilon_0(k)$ 为关联系数,由下式计算得来:

$$\varepsilon_0(k)=\frac{|\varepsilon(k)|_{\min}+\rho|\varepsilon(k)|_{\max}}{|\varepsilon(k)|+\rho|\varepsilon(k)|_{\max}} \qquad (8)$$

若对于给定的 $\varepsilon_0>0$,有 $\varepsilon>\varepsilon_0$,则称模型为关联度合格模型。

(3) 后验差检验:对残差分布的统计特性进行检验,计算数据离差 $s_1$ 及残差的离差 $s_2$:

$$s_1^2=\frac{1}{m-1}\sum_{t=1}^{m}[x^{(0)}(t)-\bar{x}^{(0)}(t)]^2 \qquad (9)$$

$$s_2^2=\frac{1}{m-1}\sum_{t=1}^{m-1}[\varepsilon^{(0)}(t)-\bar{\varepsilon}^{(0)}(t)]^2 \qquad (10)$$

后验比(均方差比值)

$$C=s_2/s_1 \qquad (11)$$

小误差概率

$$p_0=\{|\varepsilon^{(0)}(t)-\bar{\varepsilon}^{(0)}|<0.674\,5s_1\} \qquad (12)$$

若对于给定的 $C_0>0$、$p_0>0$,有 $C<C_0$,$p>p_0$,则称模型为后验合格模型。

给定一组 $\alpha$、$r_0$、$C_0$、$p_0$ 的取值,就可确定检验模型模拟精度的一个等级。常用的精度等级见表1,

**表1 精度检验等级参照表**

| 精度等级 | $\alpha$ | $r_0$ | $p_0$ | $C_0$ |
|---|---|---|---|---|
| 一级(好) | 0.01 | 0.90 | >0.95 | <0.35 |
| 二级(合格) | 0.05 | 0.80 | >0.80 | <0.5 |
| 三级(勉强) | 0.10 | 0.70 | >0.70 | <0.65 |
| 四级(不合格) | 0.20 | 0.60 | ≤0.70 | ≥0.65 |

一般情况下最常用的是相对误差检验指标。当所建立的模型残差较大,精度等级达不到要求时,或重新选取原始数列,或进行残差序列的GM(1,1)建模分析,以改善模型精度[11]。

# 2 GM(1,1)模型应用实例

## 2.1 原始数列的选取

建立GM(1,1)模型时,并非原始数列越长预测结果越好。通常情况下长序列预测的误差大于短序列。但建模数据一般都不少于5维,5维或者6维灰色预测模型的精度较高,模型拟和值与实际值最为接近[12]。当原始数列跨越有巨大差异的历史时期时,预测精度可能受到影响[5]。基于以上原因,选取了2000年人口普查后2000-2005年的我省87个县区每个县区的年末总人口数据(来源:统计年鉴)各6维作为建立预测模型的原始数列。

## 2.2 模型灰参数的计算、模型的检验

依据前一节给出的参数计算及模型检验步骤，通过编程计算得出发展系数 $a$、内生控制灰数 $b$，并进行残差、后验差的精度检验。具体实现步骤此处不再赘述。

以兰州市城关区(其它县区相同)为例，如表2—表5所示。

**表2　累加数列生成、均值生成、灰参数求解 (数列维数 $n=6$)**

| 年份 | 原始数列 $X(0)$ | 累加数列 $x(1)$ | $x(1)$的均值系列 $Z(1)$ | 灰色参数 $a,b$ |
|---|---|---|---|---|
| 2000 | 79.73 | 79.73 | | $C=1\,418.025$ |
| 2001 | 80.08 | 159.81 | 119.77 | $D=420.390$ |
| 2002 | 80.46 | 240.27 | 200.04 | $E=12\,1881.571$ |
| 2003 | 81.61 | 321.88 | 281.075 | $F=471\,194.087$ |
| 2004 | 84.26 | 406.14 | 364.01 | $a=-0.038\,49$ |
| 2005 | 93.98 | 500.12 | 453.13 | $b=73.163\,26$ |

**表3　预测数列生成**

| $\hat{X}(1)$ | $K$值 | 预测值 $\hat{X}(0)$ | 年份 |
|---|---|---|---|
| 79.730 | 0 | 79.730 | 2000 |
| 157.448 | 1 | 77.718 | 2001 |
| 238.215 | 2 | 80.767 | 2002 |
| 322.151 | 3 | 83.936 | 2003 |
| 409.380 | 4 | 87.229 | 2004 |
| 500.031 | 5 | 90.652 | 2005 |
| 594.240 | 6 | 94.209 | 2006 |
| 692.145 | 7 | 97.905 | 2007 |
| 793.891 | 8 | 101.746 | 2008 |
| 899.630 | 9 | 105.738 | 2009 |
| 1 009.517 | 10 | 109.887 | 2010 |

**表4　残差检验**

| 残差 $\varepsilon(k)$ | $\lvert\varepsilon(k)\rvert$ | 相对误差 $\Delta k$ /% | $\lvert\Delta k\rvert$/% |
|---|---|---|---|
| 0.000 | 0.000 | 0.000 | 0.000 |
| 2.362 | 2.362 | 2.950 | 2.950 |
| −0.307 | 0.307 | −0.382 | 0.382 |
| −2.326 | 2.326 | −2.850 | 2.850 |
| −2.969 | 2.969 | −3.524 | 3.524 |
| 3.328 | 3.328 | 3.541 | 3.541 |
| 平均相对误差=2.208 | | 平均精度 $p_0=0.978$ | |

通过实际计算及统计，得到兰州市城关区的人口增长GM(1,1)模型为

$$\hat{x}^{(1)}(k+1)=1\,980.5681e^{0.038\,49k}-1\,900.838\,1$$

模型平均相对误差为2.208%；方差比 $C$ 为0.456；小误差概率 $p=1$。可知该模型的精度达到了合格以上。依据该模型预测2006—2010年的人口(万人)依次为(见图1)：94.209、97.905、101.746、105.738、109.887。

依照此法逐个求得了其它各县区的GM(1,1)人口增长模型参数(略)。统计得到各个预测模型的相对误差绝对值最小值为0.000%，最大值为7.31%，有98.8%的县区相对误差最大值在5%以下(最大为4.8%)。各县区2006—2010年的具体人口预测数据本文略。

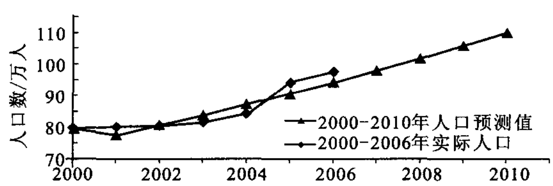


图1　兰州市城关区人口预测值

Fig. 1　Population forecasting for Chengguan District of Lanzhou city, Gansu Province.

**表5　后验差检验**

| $x^{(0)}(t)-\bar{x}$ | $(x^{(0)}(t)-\bar{x})^2$ | $\varepsilon(t)-\bar{\varepsilon}$ | $(\varepsilon(t)-\bar{\varepsilon})^2$ | $\lvert\varepsilon(t)-\bar{\varepsilon}\rvert$ | 小误差概率 $p$ |
|---|---|---|---|---|---|
| −3.623 | 13.129 | −0.015 | 0.000 | 0.015 | $<S_0$ |
| −3.273 | 10.715 | 2.348 | 5.511 | 2.348 | $<S_0$ |
| −2.893 | 8.371 | −0.322 | 0.103 | 0.322 | $<S_0$ |
| −1.743 | 3.039 | −2.341 | 5.479 | 2.341 | $<S_0$ |
| 0.907 | 0.822 | −2.984 | 8.904 | 2.984 | $<S_0$ |
| 10.627 | 112.926 | 3.314 | 10.979 | 3.314 | $<S_0$ |
| $\bar{x}=83.353$ | $\bar{\varepsilon}=0.015$ | | | | |
| $s_1=4.983$ | $s_2=2.272$ | $s_0=3.361$ | $C(s_2/s_1)=0.456$ | $p=1$ | |

## 3　结语

由于人口数据作为震害快速评估系统中重要的基础数据之一，它的精准与否直接影响到评估模型的精度。为了改善我省震害快速评估系统中人口基础数据库为静态的状况，本文从建立人口增长模型进行数据更新的角度出发，以我省87个县区2000—2005年的年末总人口数据为原始数据序列，计算得出了各个县区的GM(1,1)的人口预测模型，以便用预测得到的最新人口数据对原有静态的人口基础数据库进行及时更新，使其更接近于人口发展的实际情况，从而有助于改善震害快速评估中人员伤亡评估结果的准确程度。

## [参考文献]


[1] 杨天青,姜立新. 关于地震灾害快速评估系统的思考[J]. 地震,2005,25(3):123-127.

[2] 邹自力,刘珊红. 人口预测方法及可靠性探讨[J]. 华东地质学院学报,2002, 25 (2):142-146.

[3] 李永胜. 人口预测中的模型选择与参数认定[J]. 财经科学,2004(2):68-72.

[4] 陈彦光,余斌. 人口增长的常用数学模型及其预测方法——兼谈对 Keyfitz 双曲增长等模型的修正与发展[J].华中师范大学学报(自然科学版),2006,40(3):452-456.

[5] 郝永红,王学萌. 灰色动态模型及其在人口预测中的应用[J]. 数学的实践与认识,2002,32(5):813-820.

[6] 杨景平. 甘肃省人口规模 GM(1,1)模型及预测精度检验[J]. 开发研究,1999,(4):13-14.

[7] 黄荣清. 关于人口预测问题的思考[J]. 人口研究,2004,28(1):88-90.

[8] 王道林. 灰色预测模型 GM(1,1)及其在山东省 GDP 预测中的应用[J]. 泰山学院学报,2006,28(6):1-4.

[9] 李永胜. 人口预测中的模型选择与参数认定[J]. 财经科学,2004,(2):68-72.

[10] 郭文燕. 用灰色动态模型进行人口变动分析与预测的编程实现[J]. 市场与人口分析,2003,9(3):1-8.

[11] 刘思峰,郭天榜,党耀国,等. 灰色系统理论极其应用[M]. 北京:科学出版社,1999.

[12] 季敏,王尚旭,陈双全. 灰色理论在地球物理勘探开发中的应用综述[J]. 地球物理学进展, 2005,20(4):1164-1170.

[13] 杨天青,姜立新,杨桂岭,等. 地震人员伤亡快速评估[J].地震地磁观测与研究,2006,27(4):39－43.

[14] 马尔曼,石玉成,高晓明,等. 2003 年民乐一山丹 6.1、5.8 级地震灾害损失评估[J]. 西北地震学报,2007,29(1):64-68.

[15] 马尔曼,陈永明,赵广坤,等. 2002 年甘肃玉门 5.9 级地震灾害损失评估[J]. 西北地震学报,2004,26(2):162-167.

[16] 袁建华,于弘文,李希如,等. 从生育水平估计到未来人口预测[J]. 中国人口科学,2003,(1):15-21.

[17] 石培基,祝璇. 甘肃省人口预测与可持续发展研究[J]. 干旱区资源与环境,2007,(09):1-5.

[18] 朱伟民. Excel 在经济数学模型中的应用[J]. 微型电脑应用,2000,16(2):52-58.

[19] 门可佩,曾卫. 中国未来 50 年人口发展预测研究[J]. 数量经济技术研究,2004(3):12-17.

[20] 李强,徐桂明,范桂英. 长江中下游一南黄海地震带地震趋势研究[J]. 西北地震学报,2001,23(3):265-268.

[21] 王秀文,赵丽华. 晋冀蒙交界地区中强地震危险性的灰色拓扑预测[J]. 西北地震学报,2000,22(1):79-82.