基于时序残差辨识的终霜日灰预测模型研究

晁林海[1]，汪顺勤[2]，刘 升[3]
（1. 安徽宿州农委，安徽宿州 234000；2. 安徽宿州气象局，安徽宿州 234000；3. 安徽萧县气象局，安徽萧县 235200）

摘要　终霜冻出现的早晚是影响宿州市春播作物和越冬作物生长发育的重要因素。根据季节事件灰预测机理和时序变量灰化处理思想建立时序残差 GM（1，1）修正模型，模型精度有所提高，对终霜期的长期预测具有一定的参考价值。
关键词　终霜；时序残差；灰预测
中图分类号　S166　　**文献标识码**　A　　**文章编号**　0517－6611（2009）16－07562－02

Grey Prediction Model Research of the Last Frost Date Based on Time Series Residual Error Identification

CHAO Lin-hai et al　（Agricultural Committee of Suzhou City，Suzhou，Anhui 234000）

Abstract　The last frost appearance sooner or later is an important factor that affects spring sowing and growth of winter crop in Suzhou. According to the principle of seasonal gray prediction mechanism and the thought of time series variables gray processing，time series residual error GM（1，1）correction model is established，and model accuracy is improved，it has certain reference value for long－term forecasting the last frost time.
Key words　The last frost；Time series residual error；Gray prediction

开春之后日平均气温逐渐升高，越冬作物开始返青，春播作物亦开始下种育苗，倘若最低气温突然下降到 2℃或 0℃以下，作物生理功能就会受到破坏，这对农业的生产将产生一定影响，并随着春季终霜冻发生日期的延迟，对农业生产的影响将逐渐加重。如 4 月中下旬出现霜冻，将使处于孕穗—开花期的小麦遭到灭顶之灾，使春播作物如棉花的幼苗受冻凋萎，因此对终霜冻进行预测对合理安排农业生产具有一定的现实意义。在灰预测中终霜期的预测属于季节灾变灰预测[1]，可以直接用 GM（1，1）模型预测。但对波动较大的灾变序列直接用 GM（1，1）模型预测往往误差较大。近年来出现了许多改进方法，如干涉因子、加权模型和模型群等，这些改进模型都是基于原始数据序列进行修正的，笔者尝试把时序变量作为灰数来处理这个思路，对终霜发生日期预测模型进行修正，以提高其模型和预测精度。

1　终霜灾变序列的确定

秋季第 1 次出现的霜冻是初霜，第 2 年春季最后 1 次霜冻称为终霜，即初霜—终霜这个过程是跨年度的。笔者用宿州市 1953～2007 年终霜发生日期资料，先把终霜发生日期转换为相对测度值，相对测度序列 d 的具体求法如下：

$$d=(d(1),d(2),K,d(n))=(\boldsymbol{\delta}(1)-\boldsymbol{\sigma},\boldsymbol{\delta}(2)-\boldsymbol{\sigma},K,\boldsymbol{\delta}(n)-\boldsymbol{\sigma}) \tag{1}$$

式中，n 为序列维数（此例中为 55）；$\boldsymbol{\delta}(k)$ 为表 1 中的终霜实际发生日期；$\boldsymbol{\sigma}$ 为最早的终霜期，此例中 3 月 9 日为最早终霜期；$\boldsymbol{\delta}(k)-\boldsymbol{\sigma}$ 为相对测度值，即实际终霜期 $\boldsymbol{\delta}(k)$ 离最早终霜期 $\boldsymbol{\sigma}$ 的天数。求得宿州市 1953～2007 年终霜相对测度序列为：

$d=$（34，42，19，11，33，20，9，24，31，40，31，29，33，27，26，18，27，34，31，32，13，24，15，25，40，25，26，36，6，32，39，11，22，22，18，21，20，46，25，38，47，18，29，42，9，24，21，20，9，13，26，21，5，25，0）

以相对测度 d 作为纵坐标，年份为横坐标，作终霜事件图（F）（图 1）。

一般而言，终霜冻来得越晚对农业生产影响就越大。笔者以终霜日 4 月 15 日为例确定其灾变序列。

4 月 15 日的相对测度值为 37，在图（F）中作等高线 $d=37$，截取相对测度曲线，共得 14 个交点，每年以 10 格算，计算各交点的横坐标值，求得灾变序列 $x^{(0)}$：

$$x^{(0)}=(x^{(0)}(1),x^{(0)}(2),x^{(0)}(3),\cdots x^{(0)}(14))=(4,12,86,94,238,242,296,302,367,375,389,404,425,432) \tag{2}$$

2　预测模型的建立

2.1　时序残差修正模型建模思想　GM（1，1）模型是最常用的一种灰色模型，它是由一个只包含单变量的一阶微分方程构成的模型，是 GM（1，n）模型的特例，它是根据原始序列建立起来的，记为 $X^{(0)}$－GM（1，1）模型，建模的具体方法[2]如下：

对灾变子序列进行 1－AGO（累加生成）处理，则相应的 $X^{(0)}$－GM（1，1）模型为：

$$\frac{dx^{(1)}}{dt}+ax^{(1)}=u \tag{3}$$

利用下式求 $\hat{a}$：

$$\hat{a}=(a,u)^{T}=(B^{T}B)^{-1}B^{T}Y_{N} \tag{4}$$

把 a 和 u 带入 GM（1，1）模型，最后解得：

$$\hat{x}^{(1)}(t+1)=(x^{(0)}(1)-\frac{u}{a})e^{-at}+\frac{u}{a} \tag{5}$$

上式即为 GM（1，1）模型的时间响应函数模型。

$X^{(0)}$－GM（1，1）式是时序 t 的函数，因此给定任一时刻总能计算出 $\hat{x}^{(1)}(t)$。$X^{(0)}$－GM（1，1）模型通常存在一定误差，一是因为模型本身精度不高；二是因为与时序 t 按等时距处理有关[3]。如果假设 $X^{(0)}$－GM（1，1）模型不存在误差，即要求 $\hat{x}^{(0)}(t)=x^{(0)}$。显然，只有时序 t 存在残差 $\boldsymbol{\varepsilon}_t$，才能满足这一假设。令 $t_{\boldsymbol{\varepsilon}}=t+\boldsymbol{\varepsilon}_t$，则：

$$\hat{x}^{(0)}(t)=x\ (0)\Rightarrow\hat{x}^{(1)}(t_{\boldsymbol{\varepsilon}})=x^{(0)}(\mathrm{t})$$

$$\hat{x}^{(1)}(t_{\boldsymbol{\varepsilon}})=(x^{(0)}(1)-\frac{u}{a})e^{-a^{(t_{\varepsilon}-1)}}+\frac{u}{a} \tag{6}$$

再计算 $t_{\boldsymbol{\varepsilon}}$：

作者简介　晁林海（1969－），男，安徽砀山人，硕士研究生，农艺师，从事农作物栽培、推广和农业资源环境研究。
收稿日期　2009-03-09

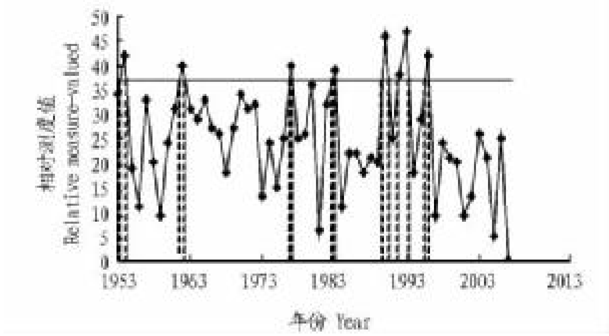

图 1　终霜事件（F）

Fig. 1　The last frost event

$$t_\varepsilon = 1 + \frac{1}{a}\ln\frac{x^{(0)}(1) - \frac{u}{a}}{x^{(1)}(t) - \frac{u}{a}} \qquad (7)$$

由 $\varepsilon_t = t_\varepsilon - t$，得时序残差序列 $\varepsilon_t^{(0)}$，对 $\varepsilon_t^{(0)}$ 进行非负处理，并建立时序残差序列 $\varepsilon_t^{(0)}$ 的 GM（1,1）模型，记为 $\varepsilon_t^{(0)}$ − GM（1,1），求得其模型值并还原得 $\hat{\varepsilon}_t^{(0)}$，设 $\hat{t}_\varepsilon(t) = t + \hat{\varepsilon}_t^{(0)}$，将 $\hat{t}_\varepsilon^{(t)}$ 替代 $X^{(0)}$ − GM（1,1）模型中的 t，求得模型值并还原，即可得到修正后的回代数据列。

预测计算时，先求出 $n+i\ (i \geqslant 1)$ 时刻的残差 $\hat{\varepsilon}_t^{(0)}(n+1)$，由 $\hat{t}_\varepsilon(n+1) = n + i + \hat{\varepsilon}_t^{(0)}(n+i)$，求预测值，即：

$$\hat{x}_\varepsilon^{(1)}[\hat{t}_\varepsilon(n+i)] = \left(x(0)(1) - \frac{u}{a}\right)e^{-a[\hat{t}_\varepsilon(n+i)-1]} + \frac{u}{a} \qquad (8)$$

还原即得最后的预测值为：

$$\hat{x}_\varepsilon^{(0)}(n+i) = \hat{x}_\varepsilon^{(1)}[\hat{t}_\varepsilon(n+i)] - \hat{x}_\varepsilon^{(1)}[\hat{t}_\varepsilon(n+i-1)] \qquad (9)$$

2.2　终霜发生日期的预测　下面以终霜 4 月 15 日为例，对其发生年份进行预测。

2.2.1　$X^{(0)}$ − GM（1,1）模型。根据（5）式求得终霜为 4 月 15 日的灾变序列（3）的 $X^{(0)}$ − GM（1,1）方程为：

$$\hat{x}^{(1)}(t+1) = 1\ 323.652e^{0.105\ 334\ 4t} - 1\ 319.652 \qquad (10)$$

式中，模型的 $a = -0.105\ 334\ 4$；$u = 139.004\ 8$

（10）式的后验差比值 $C = 0.234\ 671$，小误差概率 $P = 1$，根据文献[4]的模型精度等级划分标准，为一级模型。与原始序列之间的灰关联度为 0.635 840 1，根据经验[5]，$\rho = 0.5$ 时，关联度值 > 0.6，认为模型是“满意模型”。虽然根据文献的精度标准可以用于预测，但是（10）式的平均相对误差 $\overline{Q}$ 达 101%（$\overline{Q}$ 具体求法见公式（11）），误差较大，需要进行修正。

$$\overline{Q} = \left(\frac{\sum_{t=1}^{n}\left|\frac{x^{(0)} - \hat{x}^{(0)}(t)}{x^{(0)}}\right|}{n}\right) \times \% \qquad (11)$$

2.2.2　$\varepsilon_t^{(0)}$ − GM（1,1）模型。假设模型不存在误差，按照（7）式求得 t_ε，并由 $\varepsilon_t = t_\varepsilon - t$，得时序残差序列 $\varepsilon_t^{(0)}$

$\varepsilon_t^{(0)}$ =（0，−0.914 3，−1.321 9，−1.714 1，−1.329 4，−1.102 2，−0.789 2，−0.613 9，−0.357 5，−0.225 3，−0.178 1，−0.200 52，−0.270 35，−0.409 9）

对 $\varepsilon_t^{(0)}$ 序列数据加 2 进行非负处理得 $\varepsilon_{t1}^{(0)}$，对 $\varepsilon_{t1}^{(0)}$ 建立 $\varepsilon_t^{(0)}$ − GM（1,1）模型为：

$$\hat{\varepsilon}_{t1}^{(1)}(k) = 9.055\ 861e^{0.080\ 658\ 17(k-1)} - 7.055\ 861 \qquad (12)$$

（12）式的后验差比值 C = 0.284 341，小误差概率 p = 0.923 1，与原序列的灰关联度值为 0.612 536 6，属于“满意模型”。

2.2.3　用 $\hat{\varepsilon}_t^{(0)}$ 修正 $X^{(0)}$ − GM（1,1）模型值。对（8）模型回代值进行还原求得 $\hat{\varepsilon}_t^{(0)}$，令 $\hat{t}_\varepsilon t(t) = t + \hat{\varepsilon}_t^{(0)}$，则：

$\hat{t}_\varepsilon t1$（t）=（1，0.760 7，1.824 6，2.893 9，3.968 9，5.050 3，6.138 6，7.234 2，8.337 9，9.450 3，10.572 01，11.704 1，12.847 3，14.002 5）

用 $\hat{t}_\varepsilon(t)$ 替代（6）式中的 t，得到 $\hat{x}_\varepsilon^{(1)}(\hat{t}_\varepsilon(t))$，累减还原得 $\hat{x}_\varepsilon^{(0)}(t)$，计算方式为：

$\hat{x}_\varepsilon^{(0)}(t)$ =（4，−32.948 4，153.059 0，172.122 9，193.751 9，218.330 4，246.309 0，78.216 6，314.676 6，356.426 2，404.340 3，459.460 8，523.033 1，596.552 7）

按（11）式计算，用时序残差修正后的预测序列的平均相对误差为 49%，比 $X^{(0)}$ − GM（1,1）模型（10）的平均相对误差减少了 52%，减去第 2、3、4 个变异点，平均相对误差为 13%。由此可见，精度得到明显提高，该模型可以用于预测。

2.2.4　预测。下面对终霜日为 4 月 15 日的下一个发生年份进行预测。按（12）式求得 $k = 15$ 时的预测值 $\hat{\varepsilon}_{t1}^{(1)}$ = 20.956 0，还原再加 2 得时序残差 $\hat{\varepsilon}_t^{(0)}$ = 0.170 665。

令 $\hat{t}_\varepsilon(15) = 15 + \hat{\varepsilon}_t^{(0)} = 15.017\ 066\ 5$，用 $\hat{t}_\varepsilon(15)$ 替代（10）式中的 t，得预测值为 4 569.153，累减还原值为 681.821 5。

由于终霜灾变序列的数据是按照每年 10 格计算的，681.821 5 除以 10 约等于 68，灾变序列统计起始年为 1953，则 1 953 + 68 = 2 021，即，预测终霜日为 4 月 15 日的下一个发生年份为 2021 年，这还需有待进一步检验。

3　结论

（1）灰色系统传统 GM（1,1）模型误差较大，直接用于预测效果不理想。笔者从“假设 X 序列不存在误差而时序变量存在误差”这个角度出发，对 $X^{(0)}$ − GM（1,1）预测值进行修正，模型精度得到很大提高。对于本征性灰色系统来说，“时序残差”没有物理意义，但计算结果仍具有实用价值。

（2）预测结果从一个侧面反映了当地气候变暖和终霜期提前的事实。事实上，从 1997 年以来，终霜期均在 4 月 4 日以前，4 月中下旬没有出现过 1 次，11 年来的终霜平均日期为 3 月 25 日，比 1953～2007 年终霜平均日期 4 月 3 日提前了 9 d。

（3）以上仅对终霜日为 4 月 15 日这一个阈值进行了预测。在实际预测中，可以根据农业生产的需要，选取若干不同的终霜日期进行动态预测。可根据预测结果，采取相应的农业技术措施，减轻或避免终霜冻灾对安徽省宿州市冬小麦和棉花等作物产量的影响。

参考文献

[1] 严智渊，戴玉生. 灰色系统预测与应用[M]. 南京：江苏科学技术出版社，1989：74.

[2] 胡光宇. 战略：预测与决策[M]. 北京：清华大学出版社，2005：173 − 174.

[3] 熊和金. 徐华中. 灰色控制[M]. 北京：国防工业出版社，2005：20 − 23.

[4] 傅立. 灰色系统基础政府及其应用[M]. 北京：科学技术文献出版社，1992：68.

[5] 李华，胡奇英. 预测与决策[M]. 西安：西安电子科技大学出版社，2005：134 − 135，137.