关于GM（1，1）模型的文献综述

叶　亮
江南大学　江苏无锡　214000

摘　要：在日常做数据分析时，经常会出现这样一类数据，他们样本数量较少，信息不全，因此，在用以往的分析方法来预测和模拟这类数据时，得到的结果有时是不可用的。自灰色系统理论建立以来，灰色预测模型就解决了这一问题，即以新的方式对在一定范围内变化的、与时间有关的既包含已知又包含未知信息的灰色系统进行预测。本文主要通过对往年国内外文献进行整理，从GM（1，1）模型的发展、形式、应用、改进等方面进行了综述，并展望了其未来的发展方向。
关键词：灰色系统；GM（1，1）模型；综述

A literature review on the GM (1,1) model

Liang Ye
Jiangnan University, Wuxi, Jiangsu 214000

Abstract: In everyday data analysis, a common category of data frequently encountered is characterized by a limited sample size and incomplete information. Consequently, when utilizing traditional analysis methods to predict and simulate such data, the obtained results may at times prove unusable. Since the establishment of Grey System Theory, Grey Forecasting Models have addressed this issue by providing a novel approach to predict grey systems that encompass known and unknown information with temporal relevance and within a defined scope of variability. This paper primarily compiles domestic and international literature from previous years, presenting a comprehensive review of the GM(1,1) model, including its development, structure, applications, improvements, and further anticipates its future development trajectory.
Keywords: grey system; GM (1,1); model review

引言：

在自然和物理世界中，许多系统，如生态环境、社会、经济和科学研究，都表现为复杂和不确定的系统。这种不确定系统的主要特征在于其数据的不完整性和不充分性。信息不完整在社会、经济和科学研究活动中是普遍存在的现象，因此，可用信息的不完整性是绝对存在的，而信息的完整性则是相对的[1]。解决如何利用不完整不充分的信息，并通过建立较为完善的模型来进行未来预测，成为众多学者所探讨的重要问题。

早在古代，人们便通过观察云彩的走向来判断风向，从而预测天气情况，然而这些智慧并没有形成系统的理论体系。直到二十世纪中叶，威廉·比弗在1966年提出了单变量预测模型[2]。他通过对一定期间内的失败企业和成功企业进行比较研究，建立了各种单变量对企业财务状况进行预测的模型。随后，随着科技进步和时代发展，出现了更多适应性更强的预测模型，包括粗糙集理论、微分方程模型、回归模型等。

在1982年，北荷兰出版公司（Noah-Holland Publishing Company）的期刊《Systems and Control Letters》发表了我国学者邓聚龙教授的第一篇题为《The Control Problems of Grey systems》灰色系统论文[3]，这标志着灰色系统理论这一新兴学科的正式问世。在控制论中，人们常用颜色的深浅来形容信息的明确程度。早在1945年，控制论创始人诺伯特·维纳（Norbert Wiener）就提出了“闭盒”的概念，随后在1953年，阿什比（Ashby）首次将内部信息未知的对象称为“黑箱”[4]，这种称呼也得到了广泛接受。因此，我们用“黑”表示信息未知，用“白”表示信息完全明确，而用“灰”表示部分信息明确，部分信息不明确。相应地，信息完全明确的系统

作者简介：叶亮（2002），男，汉，安徽，本科，江南大学，江苏无锡，邮编：214000，研究方向：机器学习。

称为白色系统，信息未知的系统称为黑色系统，而部分信息明确、部分信息不明确的系统则称为灰色系统[5]。灰色系统理论中广泛应用于预测的主要模型是GM（1，1）模型，它是灰色系统理论的核心内容，也是灰色预测理论的基本模型。

GM（1，1）模型作为一种建模方法，因其所需样本数据较少、计算简便的优势，已被广泛应用于工业、农业、交通、地质、气象、生态、环境、医学、军事、经济、社会等众多科学领域。此模型在生产、生活和科学研究中成功解决了众多实际问题。特别是在面对小样本、贫信息和不确定系统的情况下，GM（1，1）模型取得了显著的应用成果，从而在预测、决策等领域中占据重要地位。灰色预测模型在应用场景和数据处理方法上，专注于研究其他模型难以解决的“小样本”和“贫信息”不确定性问题。通过充分考虑信息的覆盖，以及运用序列算子的作用，它能够揭示事物运动的现实规律。最值得强调的特点就是它能在较少的数据情况下完成有效建模，为决策者提供有价值的预测和决策依据。

一、GM（1，1）模型

1.GM（1，1）模型的原始形式

在灰色系统理论建模中，要求原始数据必须具有等时间间距。首先，通过累加生成原始数据，旨在削弱时间序列数据中的随机因素，使原本没有明显趋势的数据呈现增长趋势。再建立生成数的微分方程来进一步构建模型。GM（1，1）模型是灰色系统理论中的单序列一阶灰色微分方程，它所需信息较少，方法简便。

设$x^{(0)}$为非负的原始数据序列，$x^{(1)}$为$x^{(0)}$的一次累加生成数据序列，$z^{(1)}$为$x^{(1)}$的邻均值等权生成序列：

$$x^{(0)}=(x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n))^T,$$

$$x^{(1)}=(x^{(1)}(1),x^{(1)}(2),\cdots,x^{(1)}(n))^T,$$

$$z^{(1)}=(z^{(1)}(2),z^{(1)}(3),\cdots,z^{(1)}(n))^T,$$

其中

$$x^{(1)}(k)=\sum_{m=1}^{k}x^{(0)}(m),\quad k=1,2,\cdots,n,$$

$$z^{(1)}(k)=\alpha x^{(1)}(k)+(1-\alpha)x^{(1)}(k-1),\quad k=2,3,\cdots,n$$

$0\leq\alpha\ll1$，通常可取α=0.5，如表1

表格1　生成序列

$x^{(0)}$	$x^{(1)}$	$z^{(1)}$
$x^{(0)}(1)$	$x^{(1)}(1)=x^{(0)}(1)$	
$x^{(0)}(2)$	$x^{(1)}(2)=x^{(0)}(1)+x^{(0)}(2)$	$z^{(1)}(2)=0.5x^{(1)}(1)$
$x^{(0)}(3)$	$x^{(1)}(3)=x^{(0)}(1)+x^{(0)}(2)+x^{(0)}(3)$	$z^{(1)}(3)=0.5x^{(1)}(2)+0.5x^{(1)}(1)$
⋮	⋮	⋮
$x^{(0)}(n)$	$x^{(1)}(k)=\sum_{m=1}^{k}x^{(0)}(m)$, $k=1,2,\cdots,n$	$z^{(1)}(k)=0.5x^{(1)}(k)+0.5x^{(1)}(k-1)$, $k=2,3,\cdots,n$

以此建立灰微分方程：

$$x^{(0)}(k)+az^{(1)}(k)=b$$

称其为GM（1，1）模型，其含义为1阶（order），一个变量（variable）的灰（grey）模型（model），a称为发展系数，因为a的大小及符号反映了$x^{(0)}$和$x^{(1)}$的发展态势，b称为灰作用量或者控制系数，因为b具有灰信息覆盖作用，不能直接观测，只能通过计算得到，$z^{(1)}$称为背景值序列[6]。其中，早期在《灰色系统理论教程》中提到了a和b的显性表达式：

$$a=\frac{CD-(n-1)E}{(n-1)F-C^2},\ b=\frac{DF-CE}{(n-1)F-C^2}$$

$$C=\sum_{k=2}^{n}z^{(1)}(k)\qquad D=\sum_{k=2}^{n}x^{(0)}(k)$$

$$E=\sum_{k=2}^{n}x^{(0)}(k)z^{(1)}(k)\qquad F=\sum_{k=2}^{n}(z^{(1)}(k))^2$$

2.白化方程

通过牛顿–莱布尼茨公式和定积分我们可以得到GM（1，1）模型的白化方程：

$$\frac{dx^{(1)}(t)}{dt}=-ax^{(1)}(t)+b$$

如果我们取初始值$\hat{x}^{(1)}(t)\big|_{t=1}=x^{(0)}(1)$，可以求出其对应的解为：

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

所以

$$\hat{x}^{(1)}(k+1)=\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-ak}+\frac{b}{a};\ k=1,2,\cdots,n\quad 1$$

由于$x^{(1)}(k)=\sum_{m=1}^{k}x^{(0)}(m)$，$k=1,2,\cdots,n$，所以我们可以得到：

$$\hat{x}^{(0)}(k+1)=\hat{x}^{(1)}(k+1)-\hat{x}^{(1)}(k)=$$
$$(1-e^{a})\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-ak},k=1,2,\cdots,n-1$$

如果要对原始数据进行预测，只需要在上式取m≥n即可。

3.GM（1，1）模型评价

（1）发展系数a

模型的预测效果会因为发展系数a的取值不同而不同，刘思峰教授和邓聚龙教授（2002）[7]通过对于不同的发展系数a进行模拟分析，得到了均值模型的预测误差如表2所示

表格2　均值模型预测误差

–a	0.1	0.2	0.3	0.4	0.5	0.6	0.8	1.0	1.5	1.8
1步误差	0.129%	0.701%	1.998%	4.317%	7.988%	13.405%	31.595%	65.117%	–	–
2步误差	0.137%	0.768%	2.226%	4.865%	9.091%	15.392%	36.979%	78.113%	–	–
5步误差	0.160%	0.967%	2.912%	6.529%	12.468%	21.566%	54.491%	–	–	–
10步误差	0.855%	1.301%	4.067%	9.362%	18.330%	32.599%	88.790%	–	–	–

从而得出以下结论：

①当–a≤0.3时，GM（1，1）可以用作中长期预测；

②当0.3＜–a≤0.5时，GM（1，1）模型可以用作短期预测，中长期预测慎用；

③当0.5＜–a≤0.8时，GM（1，1）作短期预测时要慎用；

④当0.8＜–a≤1时，用残差修正GM（1，1）模型；

⑤当–a≥1时，不易采用GM（1，1）模型。

该结论也被较为广泛的应用，薛文静（2022）[8]在发展系数a=–0.06，–0.05，–0.08情况下预测了我国2020–2030年的卫生人力资源发展情况，其中通过对往年预测与实际值的比较，平均误差不超过2%，预测结果较为准确；Yang Yu（2018）在发展系数a=–0.017时较为准确的拟合了中国1990－2014年二氧化碳排放量，同时对2020–2030年排放量进行预测[9]。

（2）残差检验

为了检验预测数据与真实数据之间的差距，定义出相对误差ε（k）、平均相对误差ε（avg）和精度p^0，该检验方法也常用于其他预测模型：

$$\varepsilon(k)=\frac{x^{(0)}(k)-\hat{x}^{(0)}(k)}{x^{(0)}(k)}*100\%,$$

$$\varepsilon(avg)=\frac{1}{n-1}\sum_{k=2}^{n}|\varepsilon(k)|,$$

$$p^{0}=(1-\varepsilon(avg))*100\%$$

对于ε（k），一般要求ε（k）＜20%，此时认为模型对原数据的拟合达到一般要求，最好ε（k）＜10%，此时认为模型对原数据的拟合效果非常不错；对于p^0，一般要求p^0＞80%，最好p^0＞90%。

（3）级比偏差检验

与残差检验类似，级比偏差检验通过比较序列级比与模型级比的相对误差来检验模型的预测精度。级比偏差的计算方式为：

$$\rho(k)=1-\left(\frac{1-0.5a}{1+0.5a}\right)\frac{x^{(0)}(k-1)}{x^{(0)}(k)}*100\%$$

如果ρ（k）小于0.2，则认为可达到一般要求；如果ρ（k）＜0.1，则认为达到较高要求。

二、GM（1，1）模型的发展

1.同化阶段

在同化阶段，GM（1，1）模型尚未形成完整的体系，因此可以将其视为一般的经典数学模型，采用一般的数学概念来描述问题。在这个阶段，预测模型可以等同于一般的微分方程模型，即：

$$\frac{dx}{dt}+ax=b,\ \hat{x}^{(1)}(k+1)=\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-ak}+\frac{b}{a}$$

2.异化阶段

随着灰色理论的发展，灰色模型从微分方程模型中逐渐分离出来，此时称$x^{(0)}(k)+az^{(1)}(k)=b$为GM（1，1）模型，而在同化阶段的微分方程模型则被称为GM（1，1）模型的影子方程或白化方程，这一阶段被归属为灰色预测模型的发展阶段或同化阶段。

3.融化阶段

随着对灰色理论的不断完善和对GM（1，1）模型的深入研究，该模型进一步发展出了内涵型、派生型、离散型和延迟型等多种变体：

$$x^{(0)}(k)=\left(\frac{1-0.5a}{1+0.5a}\right)^{k-2}\frac{b-ax^{(0)}(1)}{1+0.5a},$$

$$x^{(0)}(k)=\beta-\alpha x^{(1)}(k-1),$$

$$x^{(1)}(k+1)=\beta_1 x^{(1)}(k)+\beta_{(2)},$$

$$x^{(0)}(k)+az^{(1)}(k-\tau)=bk^{\tau},$$

这些模型都具有解和方程统一的性质，并满足邓聚龙教授在《灰理论基础》一书中提出的建模的三个条件：

①结构条件：该条件要求GM（1，1）模型具有最大信息浓度的灰色导数，因为只有具备导数才能成为一个微分方程，因此这是由模型的结构决定的。

②材料条件：材料条件涉及选择哪种数据构造灰色微分方程，使得灰色导数的白化背景与灰色导数成分满足平射，这是构建方程所需考虑的材料问题。

③品质条件：品质条件要求灰色导数白化背景值位于单调增的背景集中。虽然灰色导数成分满足平射的数据可能不唯一，但只有背景值集中的数据才能确保灰色微分方程的质量[10]。

在该阶段，GM（1，1）模型经过多种变体的发展，成为了一个更加全面和灵活的预测模型。通过满足上述三个建模条件，这些模型在实际应用中能够更好地解决问题，为各个领域的预测和决策提供更可靠的支持。

三、GM（1，1）模型的应用（以经济和生态领域为例）

1.经济领域

经济系统中经济变量与未来经济发展水平指标和可持续发展指标之间的关系非常复杂，数据的完整性和充分性面临较大挑战。而且经济数据通常属于非线性时间序列，传统的预测模型往往难以准确拟合预测值与实际值之间的高精度关系，而且传统预测模型需要大量数据支持，否则预测精度可能较低。在这种情况下，GM（1，1）模型因其在有限数据下能更准确地进行预测的特性而备受关注。近年来，越来越多的学者通过灰色预测模型来对区域或相关行业的经济发展状况进行预测。李凯等（2017）和吴忠诚等（2018）分别应用灰色GM（1，1）模型对上海市2017–2020年的年度GDP和内蒙古未来四年的旅游经济总收入进行预测。这些研究表明，该模型的预测值与实际值之间的拟合程度较高，展示了其在经济预测中的应用潜力[11，12]。

2.生态环境领域

生态系统在自然中可谓是最大且最复杂的系统，而对生态环境质量进行综合评价则是实现经济社会与自然环境协调发展的重要前提。生态环境系统受多因子的综合作用，既有一定的规律性，同时又存在很强的随机性，因此被归类为典型的灰色系统[13]。在相关研究中，郑乐乐、马小雯等学者（2022）将评价模型与预测模型相结合，首先利用TOPSIS模型对指标体系进行分析研究，随后应用GM（1，1）预测模型科学地预测了甘肃省2019–2023年的生态安全指数，为当地经济、社会和生态环境的协调可持续发展提供了有价值的参考依据[14]。此外，陈媛在对社会经济与生态环境协调发展进行评价时，首先对量纲不一致的数据进行标准化处理，接着基于GM（1,1）模型分析过去时间点经济与环境的协调发展关系，并对受制当前约束下的未来区域经济、环境协调发展趋势进行了预测探讨[15]。

四、GM（1，1）模型的拓展

1.新陈代谢GM（1，1）模型

在现实生活中，大多数复杂系统往往受到周边环境和多种因素的影响。且随着时间的推移，将会不断地有一些扰动和驱动因素进入系统，使系统的发展相继地受其影响。因此，在进行长期数据预测时，使用相同的发展系数和灰作用量可能导致预测精度逐渐降低。为了解决这一问题，相关学者在原始GM（1，1）模型的基础上提出了新陈代谢GM（1，1）模型。其基本思想是认为越接近的数据对未来的影响越大，因此在处理数据时，不同的学者有各自的方法和见解。袁景凌教授（2005）在其研究中采用了最优初始条件选择算法，即对初始值进行优化处理[16]。

在求解GM（1，1）模型的白化方程时，取初始条件$\hat{x}^{(1)}(t)\big|_{t=1}=x^{(0)}(1)$，也就是默认了拟合曲线在坐标图上会经过$(1,x^{(0)}(1))$。而拟合曲线实际并不一定通过第一个数据点，袁景凌教授通过选择不同的初始条件建立多个模型，得到令误差和最小的初始条件$\hat{x}^{(1)}(t)\big|_{t=m}=x^{(0)}(m)$，以m值为初始条件建立的模型即为新的GM（1，1）模型；许秀莉（2002）在研究中采用了增加新信息、剔除旧信息的方法，以提高预测值与实际值的拟合精度[17]。在进行长期预测时，随着系统的发展，老数据的重要性逐步降低。因此，在不断补充新信息的同时，及时地剔除旧信息是至关重要的。这种建模序列能够更动态地反映系统最新的特征，实际上也是一种动态预测模型。

2.背景值改进

我们在利用公式$z^{(1)}(k)=\alpha x^{(1)}(k)+(1-\alpha)x^{(1)}(k-1)$来构造背景值时，得到的$z^{(1)}(k)$是一个平滑公式，在以月份或季度为时间间隔，或数据变化较为平缓时，该公式能达到较好的预测效果；但时间间隔过大或对于变化急剧的数据，模型则会出现较大的偏差。因此，通过对背景值改进来实现更完美的预测精度也是改进模型的一个重要方式。罗党通过改进$z^{(1)}(k)=\dfrac{x^{(1)}(k)-x^{(1)}(k-1)}{\ln x^{(1)}(k)-\ln x^{(1)}(k-1)}$来对模型进行优化[18]，使得模型更加适用于也适用于高增长指数序列建模，然而该模型不适用于发展系数较小的数据[19]，

基于此，刘乐教授（2009）[20]进一步改进 $z^{(1)}(k)=\frac{x^{(1)}(k)-x^{(1)}(k-1)}{\ln[x^{(1)}(k)-x^{(1)}(1)]-\ln[x^{(1)}(k-1)]}-\frac{x^{(1)}(1)\cdot x^{(1)}(k-1)}{x^{(0)}(k)-x^{(1)}(1)}$ 使得GM（1，1）模型的拟合和预测精度进一步提高。还有谭冠军将固定权数更改为变权数以及利用试探法和经验公式法来确定权数取值，都使得模型有了一定的优化[21]。

五、总结

GM（1，1）模型自创立以来，解决了很多其他预测模型难以解决的复杂系统，它在样本量少，信息不全的情况下实现了很好的拟合精度和预测效果，同时经过30年来的发展，已经形成了一个比较完善的理论体系，在许多领域的分析、预测中得到了广泛的应用。但是，对于模型的使用，任然还可以得到进一步的改进。

通过对往年各大期刊上对GM（1，1）模型的应用进行观察，发现绝大多数预测数据都以年份为单位，且为非负数据。虽然已有许多学者通过多方位改进模型使得预测曲线更加平滑，提高拟合精度，但对于时间间隔较小的数据，更多学者倾向于选择时间序列模型进行预测。

其次，GM（1，1）模型在应对实际问题时，不仅可以单独使用，还能与其他数据分析方法或模型结合，以实现更好的建模效果。例如，结合K-mean聚类和灰色关联分析可以更全面地分析数据；对于评价类模型，可以先使用层次分析法或TOPSIS模型计算指标权重，再利用GM（1，1）模型进行预测分析。此外，若存在多重共线性问题，可以先使用主成分分析进行降维，再进行预测。

经济预测在制定经济计划、提高经济效益以及提升管理水平方面扮演着重要角色。GM（1，1）模型作为一门理想的经济类预测模型，可广泛应用于生活基本资料、生活水平指数、建筑业、服务行业、农业等多个领域，用于对国家未来实体经济发展趋势和发展情况进行预测和分析，洞察经济的内在性质，为国家出台经济政策提供科学决策依据。

参考文献：

[1]Liu, S.; Yang, Y.; Xie, N.; Forrest, J. New progress of Grey System Theory in the new millennium. Grey Syst.Theory Appl. 2016, 6, 2 – 31.

[2]Beaver W H. Financial ratios as predictors of failure[J]., 1966, 4: 77–111

[3]Deng Ju-long Control problem of grey system[J]. System& Control Letter，1982，1(5)：288–294

[4]邓聚龙，灰色系统综述.世界科学，1983（07）：第1–5页

[5]刘思峰.灰色系统理论及其应用[M].北京：科学出版社，2021.

[6]邓聚龙.灰色系统理论教程[M].武汉：华中理工大学出版社，1990

[7]刘思峰，邓聚龙.GM（1，1）模型的适用范围.系统工程理论与实践，2000（05）：第121–124页

[8]薛文静等，基于灰色系统GM（1，1）模型的我国卫生人力资源发展及预测.中国农村卫生事业管理，2022.42（06）：第400–404页

[9]Yu Y, Deng Y, Chen F. Impact of population aging and industrial structure on CO_2 emissions and emissions trend prediction in China[J]. Atmospheric Pollution Research, 2018, 9(3): 446–454

[10]邓聚龙.灰理论基础[M].武汉：华中科技大学出版社，2002

[11]李凯与张涛.上海市2017–2020年GDP预测研究——基于改进的GM（1，1）模型.华东经济管理，2017.31（10）：第11–15页

[12]吴忠诚，朱家明与邓卓航.基于GM（1，1）对内蒙古旅游经济的研究预测.哈尔滨师范大学自然科学学报，2018.34（03）：第40–42页

[13]任成忠.大气污染灰色预测法的研究环境工程，1996（4）：35–38

[14]郑乐乐等.基于DPSIR–GM（1，1）模型的甘肃省生态安全评价与预测.生态科学，2022.41（04）：第60–69页

[15]陈媛.基于GM（1，1）模型的区域社会经济与生态环境协调发展评价——以中山市为例.环境与发展，2017.29（03）：第258–260页

[16]袁景凌，钟珞，江琼，等.新陈代谢GM（1，1）建模与应用[J].武汉理工大学学报（信息与管理工程版），2005，27（2）：168–170

[17]许秀莉，罗键.GM（1，1）模型的改进方法及其应用[J].系统工程与电子技术，2002，24（4）：61–63

[18]罗党，刘思峰，党耀国.灰色模型GM（1，1）优化[J].中国工程科学，2003，5（8）：50–53

[19]张彬，西桂权.基于背景值和边值修正的GM（1，1）模型优化[J].系统工程理论与实践，2013，33（3）：682–688

[20]刘乐，王洪国，王宝伟.基于背景值构造方法的GM（1，1）模型优化[J].统计与决策，2009（1）：153–155

[21]谭冠军，檀甲友，王加阳.灰色系统预测模型GM（1，1）背景值重构研究[J].数学的实践与认识，2015，45（15）：267—273.