分类号： 密级：

U D C ： 学号：



硕士学位论文

**我国宏观经济数据质量的评估方法研究**

**——动态评估方法及其应用**

姓 名： 赵 晓 洋

学 科 门 类： 统 计 学

指 导 教 师： 张 维 群

2013 年 03 月

**我国宏观经济数据质量的评估方法研究**

**——动态评估方法及其应用**

**学** **科：**统计学

**研 究 生签字： 指导教师签字：**

摘要

文章通过学习和借鉴国内外宏观经济数据质量检验的先进理论成果和经验，发现大多是从体制上、逻辑上定性的研究或者用单一计量模型定量的研究，从多角度定量的研究宏观经济数据质量的成果和文献较少，为此，文章基于后者采用动态评估方法对宏观经济时间序列数据质量进行研究。

首先是构建不同的动态模型，根据所研究的时间序列数据的具体特征从不同角度建立不同的动态模型并对模型进行各种检验；其次是构建一致性评估准则，分别对这两个动态模型的拟合值和预测值进行一致性评估；最后在动态模型具有良好预测效果的基础上，综合两个模型的预测值，用检验异常值准则方法分析综合预测值与实际值的差异程度，找出离群值，并对其进行相应的统计显著性检验，以此对时间序列数据质量进行评估。

最后文章运用动态评估方法对我国宏观经济数据的质量进行实证分析，以我国某地区1978-2010年间的地区生产总值和地区固定投资总额为样本数据，在分析单个时间序列数据的趋势特征和多个时间序列数据的内在关联特征的基础上，分别建立一维确定性趋势组合模型和多维VAR模型，运用动态评估方法对该地区2011年的地区生产总值和地区固定投资总额的实际观测值进行质量评估，结果显示该地区2011年的两个相应经济指标观测值是准确的。

**关键词：**统计指标；数据质量；动态评估方法；实证分析

**China's macroeconomic data quality assessment method study**

**—Dynamic assessment method and its application**

**Discipline:** Statistics **Student Signature: Supervisor Signature:**

**Abstract**

Through studying modern theoretical results and experiences in the field of the quality of statistical data, we can discovery that mostly from institutional and logic qualitative research or with a single measurement model of quantitative research, Less quantitative research the results and literature of the quality of statistical data from multiple perspectives, So that, In this paper, with the integrated dynamic assessment methods to study the quality of statistical data.

The first is to build a dynamic model, According to the specific characteristics of the

Studied time-series data from different angles to create a different dynamic model and conducted various tests on the model; Second is to build consistency assessment criteria, Conformity assessment separately to Fitting and predicted values of these two dynamic models; Finally, on the basis of the dynamic model has good predictive effect, consolidated on the predictive value of the two models, Test abnormal value criteria compares the degree of difference of the consolidated predictive value and the actual value, Identify outliers and corresponding statistical significance test, Based on that to assess the quality of time-series data.

This paper uses the prediction error analysis method based on the integrated dynamic model to empirical analysis of the quality of China's macroeconomic data. Of a regional GDP and gross fixed investment in the 1978-2010 period as the sample data, Basis on the trend characteristic sequence analysis of a single statistical indicators and the inherent associated

Characteristics of multiple statistical indicators sequence, Respectively building a one-dimensional deterministic trend portfolio model and multi-dimensional VAR model, Base on that building dynamic integrated model, Using the prediction error analysis method to

Assessment the actual observations data quality of the regional GDP and gross fixed investment of the region in 2011, The results show that the two corresponding statistical indicators observations are accurate in the region in 2011.

**Key Words:**: statistical indicators; Data quality; Dynamic integrated model; Empirical Analysis

学位论文独创性声明

本论文是我个人在导师指导下进行的研究工作及取得的研究成果。论文中除了特别加以标注和致谢的地方外，不包含其他人或其它机构已经发表或撰写过的研究成果。其他同志对本研究的启发和所做的贡献均已在论文中作了明确的声明并表示了谢意。

作者签名： 日期：

学位论文使用授权声明

本人完全了解西安财经学院有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其它复制手段保存论文。保密的论文在解密后遵守此规定。

作者签名：导师签名：日期：

目 录

[摘要](#_Toc686192060) 2

**[Abstract](#_Toc686192061)** 2

[1 绪论](#_Toc686192062) 4

[1.1 选题背景及研究意义](#_Toc686192063) 4

**[1.1.1](#_Toc686192064)** [选题背景](#_Toc686192064) 4

**[1.1.2](#_Toc686192065)** [研究意义](#_Toc686192065) 5

[1.2 研究主要内容和重点](#_Toc686192066) 5

**[1.2.1](#_Toc686192067)** [研究主要内容](#_Toc686192067) 5

**[1.2.2](#_Toc686192068)** [研究重点](#_Toc686192068) 5

[1.3 研究思路与研究方法](#_Toc686192069) 5

**[1.3.1](#_Toc686192070)** [研究思路](#_Toc686192070) 5

**[1.3.2](#_Toc686192071)** [研究方法](#_Toc686192071) 5

[1.3 可能创新和难点](#_Toc686192072) 5

**[1.3.1](#_Toc686192073)** [可能创新](#_Toc686192073) 6

**[1.3.2](#_Toc686192074)** [难点](#_Toc686192074) 6

[2 文献综述与评价](#_Toc686192075) 6

[2.1 统计数据质量评价的国外文献综述](#_Toc686192076) 6

[2.2 统计数据质量评价的国内文献综述](#_Toc686192077) 6

**[2.2.1](#_Toc686192078)** [从计量模型的角度对数据质量进行评估](#_Toc686192078) 6

**[2.2.2](#_Toc686192079)** [从异常值的角度对数据质量进行评估](#_Toc686192079) 6

**[2.2.3](#_Toc686192080)** [其它统计数据质量评估方法](#_Toc686192080) 7

[2.3 文献评价](#_Toc686192081) 7

[3 动态评估方法的基本思想](#_Toc686192082) 7

[3.1 基本假设](#_Toc686192083) 7

[3.2 动态模型的建立](#_Toc686192084) 7

**[3.2.1](#_Toc686192085)** [一维动态模型](#_Toc686192085) 8

**[3.2.2](#_Toc686192086)** [多维动态模型](#_Toc686192086) 8

**[3.2.3](#_Toc686192087)** [模型预测准确性评价](#_Toc686192087) 8

[3.3 一致性评估准则](#_Toc686192088) 8

**[3.3.1](#_Toc686192089)** [距离法](#_Toc686192089) 9

**[3.3.2](#_Toc686192090)** [变异系数法](#_Toc686192090) 9

**[3.3.3](#_Toc686192091)** [一致性相关系数法](#_Toc686192091) 10

[3.4 数据质量评估](#_Toc686192092) 11

**[3.4.1](#_Toc686192093)** [离群值评估方法](#_Toc686192093) 11

**[3.4.2](#_Toc686192094)** [异常值显著性评估准则](#_Toc686192094) 13

[4 动态评估方法的实证分析](#_Toc686192095) 15

[4.1 基本情况说明](#_Toc686192096) 15

**[4.1.1](#_Toc686192097)** [研究变量](#_Toc686192097) 15

**[4.1.2](#_Toc686192098)** [数据说明](#_Toc686192098) 15

**[4.1.3](#_Toc686192099)** [基本假设](#_Toc686192099) 15

[4.2 一维动态模型的建立](#_Toc686192100) 16

**[4.2.1](#_Toc686192101)** [地区固定资产投资总额序列一维动态模型的建立](#_Toc686192101) 16

**[4.2.2](#_Toc686192102)** [地区生产总值序列一维动态模型的建立](#_Toc686192102) 18

[4.3 多维动态模型的建立](#_Toc686192103) 20

**[4.3.1](#_Toc686192104)****[VAR](#_Toc686192104)**[模型构建](#_Toc686192104) 20

**[4.3.2](#_Toc686192105)****[VAR](#_Toc686192105)**[模型预测效果检验](#_Toc686192105) 22

[4.4 一维和多维动态模型一致性评估](#_Toc686192106) 23

[4.5 数据质量评估](#_Toc686192107) 30

[4.7 可知，lntz1 的误差百分比绝对值为|1](#_Toc686192108) 31

[5 结论与展望](#_Toc686192109) 31

[5.1 结论](#_Toc686192110) 31

[5.2 展望](#_Toc686192111) 32

[参考文献](#_Toc686192112) 32

[攻读硕士学位期间发表的论文](#_Toc686192113) 33

[附录](#_Toc686192114) 34

# 1 绪论

## 1.1 选题背景及研究意义

### **1.1.1** 选题背景

近年来，我国政府统计系统体质改革力度逐步加强，大规模统计数据搜集日趋系统化、制度化、国际化，然而由于统计工作系统复杂庞大，以及统计普及工作在我国还处于初级阶段，我国不管是宏观还是微观方面统计数据质量都不容乐观，国内外学者对我国的统计数据尤其是宏观经济统计数据的质疑愈演愈烈。与此同时，确保统计数据质量的难度也在不断加大。

2012年全国各省（区、市）核算出的地区生产总值相加总量达到57.69万亿元，比国

家统计局此前公布的2012年初步核算的国内生产总值51.93万亿元高出5.76万亿元，相当于多出一个广东的经济总量。这种现象已经持续多年。

据介绍，自1985年国家和地方层面分别核算地区生产总值数据以来，地方统计总和一直高于全国的GDP总量，不仅呈现出“1+1> 2”的局面，而且有递增的趋势。例如，2009年各省地区生产总值之和超出全国2.68万亿元；2010年各省地区生产总值之和超出全国

3.2万亿元；2010年31省区市地区生产总值超出全国总量3.5万亿元；2011年31省区市地区生产总值总和超出全国总量4.6万亿元，而2012年地方地区生产总值之和竟然超出全国5.76万亿元之多。

国外学者对我国宏观经济统计数据的质量的质疑声一直不断，甚至有外国学者认为：中国经济数据，是一对不可置信的符号体系。看中国的经济统计数据，就好像在看一部科幻小说。在1988年Perkins就怀疑中国经济增长率存在上偏误差[1]，2001年美国Rawski连续发表两篇文章认为中国官方公布的经济增长数据几乎都是不可信的[2][3]。

2012年3月，国家统计局还曝光了地方政府干预企业统计数据上报造假的案例，国家统计局局长马建堂多次强调惩治和预防统计弄虚作假行为。但是，地方地区生产总值总

和高于全国数据的问题长期得不到解决，而且呈现递增趋势，不知是统计技术差异，还是执法力度不够？

有的地方在统计地区生产总值时搞双轨制：“摆政绩时是一个数字，要补贴时又是一个数字”，甚至身兼“百强县”、“贫困县”两个头衔。如果没有扎实的经济发展与民生的改善，表面上的地区生产总值繁荣，非但没有意义，还可能误导宏观决策，挤占民生空间。而有的地方官员还躺在地区生产总值政绩的沙发上沾沾自喜。统计数字注水、地区生产总值层层加码，地区之间互相攀比的现象屡见不鲜。

为了提高我国统计数据的质量，国家统计局局长马建堂在中国统计学会第八届一次常务理事会上的讲话中也指出：要深入落实科学发展观，紧紧围绕提高统计能力，提高数据质量和公信力，一是要着力强化对提高统计能力的研究，二是着力强化对提高统计数据质量的研究，三是着力强化对提高统计公信力的研究。这就要求统计工作者在用现代科学方法对如何探索和解决统计数据质量的问题方面加大精力。

### **1.1.2** 研究意义

统计工作好比统计信息的加工生产活动，而统计数据就是统计工作的核心材料，因此，统计数据质量就是是统计工作的基石和生命线，统计数据质量的好坏直接决定统计工作的成败，进而关系到以统计数据为依据做决策的个人及政府、企业的利益损失。随着不同的用户、生产者和被调查者三个角度来看待统计数据，统计数据质量由狭义的单方面的准确性发展为广义的，具体包括准确性、实用性、可衔接性、及时性、可比性、有效性、可解释性、可取得性、客观性（或诚信）、方法专业性（或健全性）、减轻调查负担等11个方面的统计数据质量标准[4]。不同的统计数据用户、生产者和被调查者在不同的时期对统计数据质量有不同的标准，这又使得对统计数据质量的诊断准则不一，针对不同的统计数据质量标准构造相应的诊断准则是迫切的也是必要的。

目前我国对宏观经济数据质量问题的研究很多，但大多是从体制上和逻辑上用定性的的研究方法来研究宏观经济的数据质量，定量的相关研究较少。在用计量模型研究宏观经济数据质量也有相关的文献，但是基本上都是从一个角度用单一的计量经济模型来研究统计数据的质量问题，从多角度综合多个模型的结果来研究该问题的文献几乎没有，为此，本文采用多角度建立综合模型方法进行统计数据质量评估。

文章运用动态综合模型的预测误差分析方法对统计数据质量（狭义上的准确性标准）进行评估，根据每个指标序列的趋势特征和多个指标序列的内在关联特征，分别从一维和多维的角度建立动态模型，在两个动态模型通过各种检验和预测效果一致性检验，构建动态综合模型，在其具有良好的预测效果基础上，运用检验异常值方法通过比较预测值和实际观测值之间的差异程度，找出离群观测值，再进行相应的统计显著性检验，对我国宏观经济数据的质量进行综合评估。与单一模型相比，动态综合模型在一定程度上能提高对统计数据质量评估的精度，可以对我国统计数据质量可靠性的评估起到一定的参考作用。

## 1.2 研究主要内容和重点

### **1.2.1** 研究主要内容

文章以我国宏观经济时间序列数据为研究对象，运用动态评估方法对其数据质量的准确性进行评估研究。主要介绍动态评估方法的基本思想和实证分析。

#### 1）动态评估方法的基本思想

首先是构建不同的动态模型，根据所研究的宏观经济时间序列数据的具体特征从不同角度建立不同的动态模型并对模型进行各种检验；其次是构建一致性评估准则，分别用所构建的动态模型对相关的时间序列数据进行拟合，再对这两组拟合值进行一致性评估；最后在动态模型具有良好预测效果的基础上，加权综合两个模型的预测值，用检验异常值准则方法分析综合预测值与实际值的差异程度，找出离群值，并对其进行相应的统计显著性检验，以此对时间序列数据质量进行评估。

#### 2）实证分析

以我国某地区1978-2010年间的地区生产总值和地区固定投资总额为样本数据，在分析单个时间序列数据的趋势特征和多个时间序列数据的内在关联特征的基础上，分别建立一维确定性趋势组合模型和多维VAR模型，运用动态评估方法对该地区2011年的地区生产总值和地区固定投资总额的实际观测值进行质量评估，结果显示该地区2011年的两个相应经济指标观测值是准确的。

注：若不特殊说明，后文中所提及“时间序列、统计数据”等均是指“宏观经济时间序列数据”。

### **1.2.2** 研究重点

主要研究动态评估方法的两方面的评估准则：一是不同动态模型拟合效果和预测结果的一致性评估准则；二是宏观经济时间序列数据质量的预测误差评估准则，包括离群值评估准则和异常值显著性检验准则。

## 1.3 研究思路与研究方法

### **1.3.1** 研究思路

由上文“研究重点”中可知，宏观经济时间序列数据动态评估方法主要研究思路是从它的两方面准则展开，具体思路流程如下：

动态评估方法关键在于三个检验评估一是不同动态模型拟合效果和预测结果的一致性评估；二是基于综合预测值的预测值误差离群值评估；三是异常值显著性检验。

具体研究思路如图1.1所示：



动态模型 1

动态模型 2

N

一致性评估

Y

N

预测误差离群值评估

Y

N

异常值显著性检验

Y

数据质量审查及处理

时间序列数据可疑

综合预测结果

时间序列数据准确

时间序列可靠、离群

时间序列数据异常

图1.1 研究思路示意图

### **1.3.2** 研究方法

本文综合运用了经济学、计量经济学、时间序列分析和统计学等相关理论运用动态分析方法进行统计数据质量的评估。具体运用以下研究方法：

一是规范分析和实证分析相结合。文章基于相关理论，运用计量动态分析方法和误差分析方法构建时间序列数据质量的评估方法，同时以我国某区域宏观经济数据进行实证研究。

二是一维和多维分析相结合。本文通过分别建立一维动态模型和多维动态模型，分别刻画出单个时间序列的趋势特征和多个时间序列的内在关联规律，综合一维和多维动态模型的预测结果对时间序列数据质量进行评估。

三是基于动态模型的异常数值识别分析方法。文章先通过两个动态模型模型分别刻画出统计指标之间的内在联系和每个统计指标的趋势，运用异常值检验准则检验模型的预测值和实际观察值的差异程度，找出离群值，并进行相应的异常值显著性检验，以此对时间序列数据质量进行评估。

## 1.3 可能创新和难点

### **1.3.1** 可能创新

1）从一维和多维视角分别对单一统计指标序列和多个统计指标序列建立动态模型；

2）推广“计量模型的预测误差分析方法”到“动态综合模型的预测误差分析方法”，即由单一模型推广到多个模型，综合多个模型的预测结果进行统计数据异常值评估；

3）借用衡量两个系统协调程度的协调系数——变异系数法——评估两个不同动态模型性能的一致性程度。

### **1.3.2** 难点

#### 1）样本容量可能不足。多维动态模型的一个特点是参数较多，如VAR模型，一个VAR

模型含有三个变量，最大滞后期*k* = 3，则有*k N* 2 +N= 332 +3= 30个参数需要估计。当

样本容量较小时，多数参数的估计量误差较大。如果取年度资料，从1978-2010年共33

个，样本量不足；若干取季度或月度宏观经济统计资料有时则很难获取。

#### 2）历史资料假设性较强。鉴于计量动态模型的特性——假设历史数据是准确的，这个假设在现实中往往不能得到满足，导致模型预测的结果不准确，从而对相关时间序列数据质量的可靠性评估失真，如何确保历史数据的准确性比较困难。

#### 3）部分时间序列数据从两个或者两个以上角度进行研究比较困难。一般从一个角度较好地对某些事物进行研究尚且困难，更何况从多个不同角度都较好的对其进行研究，因此，用动态评估方法某些时间序列数据质量进行评估实施起来较困难。

# 2 文献综述与评价

## 2.1 统计数据质量评价的国外文献综述

国外有关统计数据质量的研究始于20世纪初，研究至今大致可分为三个阶段：第一阶段从20世纪初至40年代，国外学者主要研究对象是统计数据的准确性和样本的代表性，具体分为两个研究方向。一是对统计数据的准确性和抽样误差的研究。1934年著名统计学家奈曼(C. J. Neyman)的研究成果形成了抽样方法中的“可测设计”，这对用统计理论方法统计数据质量研究有着里程碑的意义。二是统计数据质量的准确性和非抽样误差的研究。1915年，英国统计学家鲍莱(A. L. Bowley)首次提出了非抽样误差这一概念，政府统计的大规模应用使得对非抽样误差的研究不断深入，实际上就是研究非抽样误差对统计数据质量的影响。第二阶段从上世纪50年代初至70年代，该阶段主要是构建统计调查误差模型对统计数据质量进行研究。突出研究成果为Hansen和Hurwitz（1946）提出的无回答问题这一概念[5]，其后Hansen, Hurwitz和Bershad（1961）首次提出了完整的调查误差模型[6]；20世纪60年代，许多学者开始研究均方误差的主要组成部分，并试图减少这些成分，如Warner（1965）提出了对敏感性问题调查的随机化方法[7]。第三阶段从20世纪70年代之后，该阶段统计数据质量的研究日趋成熟，随着统计数据质量涵义的夸大与完善，开始研究如何建立有效的保证、控制和评估体系这一层面。联合国统计局与1980年出版的《统计组织手册》中，明确对官方统计数据提出了八项要求和政府统计数据质量的综合涵义[8]。在统计数据质量涵义的演变进程中，Dalenius（1983）提出统计数据质量的“测量向量”，具体分为统计数据的经济性、时效性、准确性、相关性、保密性和详细程度等方面[9]。Brackstone（1999, 2000）提出统计数据质量的六个维度——准确性、一致性、相关性、可解释性、及时性和可取得性[10]。Frschl和Grossman（2000）提出了全面的统计数据质量管理的统计、执行框架，将统计数据的各个维度与统计数据生产过程一一配对[11]。此外，20世纪90年代中期国际货币基金组织(IMF)分别建立了数据公布特殊标准(SDDS)和数据公布通用系统(GDDS)以及数据质量评估框架(DQAF)等国际标准，对统计数据的产生、公布以及数据质量定性评估提供了标准的指导方法。在国际统计数据质量研究不断完善的同时，由不断产生新的统计数据质量评估方法，如逻辑分析法、基于模型分析法等，其中比较有代表性的有：Rawski (2001)[2]运用逻辑分析法通过分析我国1998-2001年间

的官方公布数据，指出中国经济增长率与消费与收入、能源消耗、生产与投资等统计数据之间都存在不一致性。Rawski得出这一结论的逻辑前提是假设GDP与能源消耗等应该保持一致增长，而这一前提仅仅是主观定论，且从其他发达国家如美国、英国、日本和韩国等的发展实践来看，值得商定。Klein、Ozmucur（2002）分析所研究的统计指标序列特征构建拟合模型，对所拟合模型中的各个解释变量的系数进行深入分析，判断拟合模型所刻画的被解释变量与解释变量之间的内在规律是否与社会经济常识或构建模型时的预期一致的角度对统计数据质量的可靠性进行评估[12]。例如他们在研究国内生产总值或其增长率的可信度时选取了来源相对独立的、尽可能多的和代表性强的经济指标作为解释变量，为排除多重共线性的影响，同时也为节省自由度，他们在估计模型参数前先对原始自变量做了主成分分析再将得到的主成分因子作为自变量进行参数估计，并根据主成分与原始指标的关系将各主成分因子的估计系数还原成原始指标的系数，该文献不仅在理论上的阐述了运用计量模型方法进行统计数据质量评估的可行性，还运用实证分析详述了计量模型方法进行统计数据质量评估的操作流程，对运用计量模型方法探讨统计数据质量的问题有里程碑意义。

## 2.2 统计数据质量评价的国内文献综述

### **2.2.1** 从计量模型的角度对数据质量进行评估

所谓计量模型分析法是以构建计量经济模型为基础，刻画出所研究的统计指标序列的趋势特征或者指标之间的内在关联特征，对统计指标的数据质量进行诊断的一种统计数据质量评估方法。目前，用于统计数据质量诊断的计量模型大致可分为以下四类：传统回归模型、经典时序模型、面板数据模型和其它计量模型，后者包括自回归分布滞后模型、误差修正模型、曲线回归模型、联立方程模型以及向量自回归模型等。传统回归模型凭借成熟的经济理论分析能有效的刻画出经济指标之间内在的关联规律，阙里和仲笑寒（2005）利用我国28个地区1984-2001年间的10个主要经济量，一共504个样本数据，运用主成分分析和固定效应变截距模型对我国国内生产总值的增长率的准确性进行了地区层次的评估，结果表明不能证明我国国内生产总值数据存在系统的、长期的错误[13]。刘孝新等

（1997）使用事后预测与反常结果判断法对统计数据质量的准确性进行评估，提供了一套完整的评估统计数据质量准确性的研究方法，但是该方法仅从变量之间的经济理论关系出发建立多变量计量模型，其假设前提“全部统计数据来自同一总体在相同或相近的经济系统环境下的统计数据”是不合理的，缺少对统计指标时间序列波动情况的分析[14]。经典时序模型依据对统计指标历史数据的深入分析能有效的刻画出所研究统计指标数据的变动规律。周潮（2008）使用季节时间序列模型(Seasonal ARIMA Model, SARIMA)，根据

统计指标历史数据的趋势特征建立模型，在模型通过各种检验（比如统计、经济、计量等相关检验）和具有良好预测效果的基础上，通过分析模型拟合出的预测值与实际值的差异程度，找出离群样本值，并进行相应的统计显著性分析，从而能有效的对统计指标的数据质量进行诊断[15]。面板数据模型不仅能有效的刻画出宏观经济现象之间复杂的内在规律，还能描述出上述复杂的内在规律在不同时期或不同个体之间的差异；其它经济模型如从静态扩展到动态，从一维扩展到多维，例如杨海ft、许启发（2001）选取我国某地区1978～

1995年间的地区生产总值和社会消费品零售总额这两个相关性很强的统计指标作为研究对象，利用向量自回归模型（VAR模型）对该地区1996～1998年间的地区生产总值的准确性进行评估[16]。刘洪和黄燕（2007, 2009）先后运用趋势模拟评估法和相关指标分析法，对统计数据质量进行了定量评估[17][18]。张维群等（2010）统计数据质量定量诊断与控制方法研究，从静态和动态分别建立区域宏观经济统计数据质量诊断模型，综合两个模型的结果进行统计数据质量的诊断[19]，这种从多视角评估统计数据质量的方法在国内研究较少。

计量模型方法能刻画经济系统中多经济指标间的复杂规律，但是这种方法对假设条件要求较苛刻，模型的参数估计对估计方法较敏感，因此基于模型参数的统计数据质量评估方法要对统计数据进行必要的处理或选用比较稳健的参数估计方法，如赵进文（2010）通过列举异常值对序列相关性检验、多重共线性关系检验、单位根检验和异方差性检验等经济计量检验结果产生致命影响的典型案例，说明异常值处理对经济计量建模的重要性[20]。卢二坡、黄炳艺（2010）首次使用基于稳健MM估计的异常值评估方法，分别使用两种不同的劳动力投入数据对所构建的生产函数进行拟合，评估我国改革开放以来国内生产总值的数据质量，该方法可有效地解决传统方法中经常出现的多个异常点掩盖的弊端[21]。卢二坡、张焕明（2011）首次使用稳健的主成分回归方法及异常值评估方法，对2008年我国地区经济增长横截面统计数据的可靠性进行了诊断。结果表明，异常值对稳健的主成分回归方法的影响很小，使评估结果更加可靠，并能有效的克服传统主成分回归方法容易被多个异常点的掩盖现象[22]。

### **2.2.2** 从异常值的角度对数据质量进行评估

异常值评估分为单变量和多变量异常值检验，又可分为参数和非参数异常值检验。单变量异常值参数检验假定给定的统计资料服从一个随机分布，并用不一致性检验来识别异常点。在假设条件资料服从正态分布的前提下，传统的的异常值评估方法，如3准则、

Grubbs准则、Dixon准则、Chauvenet准则、Nari准则和偏度—峰度检验等，3在大样本

情况下使用是最简便的，但是在小样本时是失效的；Grubbs准则使用于剔除一个异常值；

Dixon准则使用于小样本数据中的异常值检验，能一次剔除多个异常值，比Grubbs准则更简便，但是这些准则都受极端异常值的影响；Chauvenet准则一般使用于小样本数据；

Nari准则要求方差已；这些传统的异常值评估准则稳健性都较差且对检验多个异常值（尤其是在同侧）的效果很差，在样本量较大，且异常值个数不很明确时评估异常值，发展出

2 文献综述与评价

稳健统计法，如楼润瑜（2008）Robust Z检验法[23]和朱宏（1989）样本分位数法[24]，Robust

Z检验法将极端值给统计参数带来的影响减至最低，使用中间值代替均值，无需将极端值从测量数据中删除，而是减低其权重；样本分位数法使用样本分位数的优化组合来估计总体参数，在正态样本的异常值检验中具有较好的稳健性，和前面传统的方法相比，有计算简便、抵抗异常值污染能力强和可连续使用检验多个异常值等优点。异常值检验的非参数方法常用的有五数概括法和Walsh法[25]，五数概括法操作简易，对大样本检验功效较高，但对小样本则功效较差；Walsh检验法适用于任意分布对象，可同时检验多个可疑数据，但它不适用于大样本情形。成邦文等（2000、2003）研究发现：一维数据与多维数据的宏观经济统计指标近似服从对数正态分布，因而提出统计数据质量诊断和异常值评估的对数正态分布置信区间检验法，并且从一致性误差和非一致性误差两个方面分析了宏观经济统计数据误差，关于前者的分析内容有机构对比分析、趋势性分析、和相关经济指标关联分析等；而对后者的分析则包括统计指标数据分布形态一致性检验和异常数据的评估[26][27]。张德然（2003）给出了分析各统计指标数据点的跳跃度变化程度来评估异常值的方法，该方法不仅能处理有异常小值或异常大值的异常值评估方法还能诊断既有异常小和异常大值的检验，弥补了以往异常值检验方法只能检验前部分的情况，但是检验的异常值仅仅是指狭义的离群值，并且还没有用相关统计量进行离群值显著性检验，对异常值的评估效果主观性较强，可能会因离群值定义的标准不同而得出不一致的评估标准[28]。李子奈和周建

（2005）从经济系统角度出发，使用联合估计诊断方法对我国宏观经济指标序列进行探索分析，结果表明，联合估计诊断方法能有效的刻画出统计指标序列中异常值的复杂现象[29]。刘洪、黄燕（2009）在文中采用学生化残差、Cook统计量、W-K统计量作为异常值的检测统计量对我国相关统计资料进行了实证分析[18]。上述主要是针对单个变量的异常性进行诊断，若一般若是同时检验多个变量的异常性时，异常值检验要复杂的多，不仅要对每个单独变量进行异常值检验，还要对多个变量之间的协调性进行检验，进一步判断其异常性。在样本量较大时，多变量异常性评估的一般方法有，回归诊断法、杠杆值检测法、马氏距离(MD)检测法和“投影寻踪”法等。回归诊断法严重依赖模型假设、极端异常值也对其参数估计影响很大，杠杆值检测法需要模型满足经典假设条件，马氏距离(MD)检测法由于期望和协方差都的估计量不稳健，使得马氏距离(MD)检测具有不稳健性，因此，这三种方法对多变量异常值检测效果都较差，为了解决该弊端，发展出一种基于稳健马氏距离

（RD）异常值检测方法，该方法主要基于快速MCD方法使用迭代和一般马氏距离的思想构造稳健的协方差和稳健的均值向量估计量，得到一个稳健的多变量异常值检测方法[30]。

“投影寻踪”法是一种对多维数据的异常性诊断比较有效的一种方法，朱文杰（2007）证明了“投影寻踪”法既能诊断出多维数据中的异常值点又能充分保留和使用多维数据的关联性和结构性[31]。

西安财经学院硕士学位论文

### **2.2.3** 其它统计数据质量评估方法

其他统计数据质量评估方法，在统计方面主要从体制、逻辑、误差、贝叶斯等角度研究统计数据的质量诊断问题，这些统计数据质量评估方法，主要是定性的研究分析，在技术性研究方面比较少，对于数据质量的技术诊断缺乏可操作性[32-43]；在非统计方面，则主要从数据库、博弈论、数据挖掘等角度研究，这些数据质量评估方法主要研究对象是大数据，对小数据的质量评估效果较差，且这些数据质量评估方法大多数没有进行相应的统计显著性评估，导致主观性和随机性对统计数据的质量评估效果产生一定的不确定影响

[44-50]。

## 2.3 文献评价

从国内外学者对于统计数据质量的研究中可以看出，国外主要研究广义的统计数据质量，至于统计数据质量的定量诊断主要采用单个统计指标的定量诊断，缺乏宏观经济诸多指标的系统性诊断；国内关于统计数据质量的研究重点在于狭义概念上的统计数据质量，多集中于或定性分析、或立足于数据质量问题的成因、或从政策角度来分析、或静态定量分析，对变量尤其是多变量的动态方法分析研究较少。

鉴于此，本文以统计指标中时间序列数据的动态性和系统性相结合，运用动态评估方法对时间序列数据质量进行评估，即通过时间序列之间的内在关联规律建立多维动态模型；通过单一时间序列的趋势特征建立相应的一维动态模型，通过对两个模型进行各种检验，具有良好预测效果的基础上，综合两个动态模型的预测值，运用预测误差评价方法对时间序列数据质量进行评估。

# 3 动态评估方法的基本思想

当宏观经济时间序列数据有异常时，通过分析时间序列数据的趋势特征所构建的一维动态模型因含有异常趋势的信息而不能真实刻画出相应时间序列数据的内在趋势规律，这时该一维模型的预测效果的准确性就大打折扣；同理，当多个宏观经济时间序列的历史数据都发生同样的异常性，这时所构建的多维动态模型即使拟合效果再好，因包含时间序列之间的异常信息，并不能反应其真实的内在关联规律，其预测效果的准确性同样大打折扣。上述表明，当时间序列数据存在异常时，不论单独从哪个角度构架动态模型所刻画的时间序列的内在规律都不能排除异常数据的影响，为了削减这种弊端，可以采用同时从多个角度研究时间序列的内在规律来构建动态模型，通过不同模型的拟合效果和预测结果的一致性评估可以诊断样本数据是否异常：

#### 1）当一致性评估通过时，说明所研究时间序列数据质量可靠；

#### 2）当一致性评估不通过时，说明所研究时间序列数据质量可疑。

因为只有当事物本身是正常发展的情况下，从不同角度都能预测其未来的发展变化，且预测的结果是一致的殊途同归的；相反，当事物本身某时期发展异常时，从不同角度所反映的预测结果是不一致的。因此，可以通过预测结果是否一致来推测该时期事物发展是否正常，用在动态模型上就是对其进行拟合效果和预测结果的一致性评估。

因此，为了降低单一模型的这种弊端，提高动态模型法对时间序列数据质量评估的效果，文章采用动态评方法对时间序列数据质量进行评估。

动态评估方法的基本思想是指对宏观经济时间序列数据分别从不同的角度建立动态模型，根据每个时间序列数据的趋势特征建立一维动态模型，对时间序列的变化趋势进行刻画；根据多个时间序列之间的内在关联规律构建多维动态模型，对时间序列之间的复杂内在关系进行刻画，在两个模型拟合效果通过各种检验的基础上，综合两个模型的预测值，通过比较实际观测值和综合预测值来对相关宏观经济时间序列数据质量进行评估。

## 3.1 基本假设

#### 1）假定大部分宏观经济时间序列历史数据是真实可靠的。

我国宏观经济指标数据来自于国家和地方统计局等权威机构，因此假定大部分宏观

经济指标数据时真实可靠的是合理的。

#### 2）任何一个经济指标时间序列都存在着特定的趋势规律，该规律可以通过构建一维动态模型刻画。

任何一个经济指标在宏观经济体系中都有其特有的趋势规律，其发展过程就是一个时间序列过程，因此经济指标的发展变化能够通过时间序列模型刻画其发展规律。

#### 3）所选择的多个经济指标之间存在内在关联关系，该关系可以通过多维动态模型刻画。由于所选择的两个经济指标来自于同一个宏观经济系统中，同一系统中的各个因素都

是相互影响和相互制约的，这种动态发展的复杂关系符合一定客观规律，因此各经济指标之间的这种内在关系能通过多维动态模型给予刻画。

## 3.2 动态模型的建立

### **3.2.1** 一维动态模型

可以通过构建一维动态模型来刻画某时间序列数据的趋势特征。在一维动态模型的应用中，根据所研究的时间序列数据具体趋势特征构建相应的一维动态模型，具体情况如下：

当所研究时间序列数据无异方差性或者异方差性较小，可以通过对相应时间序列数据进行数据变换处理，可以采用确定性分析模型（趋势拟合模型、季节效应分析模型、综合分析模型及X-11模型等）、随机性分析模型（ARIMA等）或者确定性和随机性相结合分析的组合模型（残差自回归模型等）等进行分析；

当所研究时间序列数据含有显著的异方差性，且异方差函数不知时，这时根据实际情况可以采用条件异方差模型（ARCH）及其扩展模型（GARCH 模型、EGARCH模型、

IGARCH模型、TARCH模型、GARCH-M模型和AR-GARCH模型等）直接对相关时间序列数据进行建模分析。

实际分析中，应使用多种分析方法对所研究的时间序列数据的趋势特征进行模型分析，选择最优的一维动态模型。

### **3.2.2** 多维动态模型

根据上文“基本假设”可知，多个时间序列数据之间复杂的内在规律可以通过构建多维动态模型进行刻画。常用的多维动态模型有：传统的计量模型方法、非结构性模型（如向量自回归模型（VAR）、向量误差修正模型（VEC）等）和结构VAR模型（SVAR）等。

传统的计量模型方法是以经济理论为基础来描述经济时间序列指标之间内在关系的模型，如联立方程模型等。但是经济理论一般对经济时间序列指标之间的复杂的内在关联规律刻画的不严紧，而且内生变量可以同时出现在联立方程的两端，这使得参数估计和结

果分析变得更加繁杂。向量自回归模型（VAR模型）是基于时间序列数据的统计性质建立模型，VAR模型采用方程组的形式，在模型的每一个方程中，内生变量对模型的全部内生变量的滞后项进行回归，从而估计全部内生变量的动态关系。VAR模型在刻画多个相关经济指标序列的复杂关系和预测方面是非常有效便利的[51]。

*VAR*（*p*）模型的数学表达式是：

*Yt*1 *yt*1... *p yt**p**Hxt* *t*

*T*1, 2,..., *T*

式中：*yt*是*k*维内生变量列向量，*xt*是*d*维外生变量列向量，*p*是滞后阶数，*T*是样本个数。*k**k*维矩阵1，..., *p* 和*k**d*维矩阵，*H*是待估计的系数矩阵。*t* 是*k*维扰动列向量。

VAR模型中不包含外生变量的非限制变量自回归模型的形式如下：

*Yt*1 *yt*1... *p yt**p* *t*

由VAR模型的结构可知VAR有以下优点：

*T*1, 2,..., *T*

1）不以经济理论为基础。建模简易，只需完成两件事：（1）找出有相互关系的变量，使之包含在VAR模型中；（2）确定最大滞后期数，使变量间的相互关系在VAR模型得到最大的体现；

#### 2）VAR模型对参数没有零约束，即参数估计值显著性与否都保留；

#### 3）VAR模型的自变量中不包括任何当期变量，有效解决联立方程模型有关的弊端。

4）无约束VAR模型有良好预测性能。由于在VAR模型中每个方程的右侧都不含有当期变量，这种模型用于预测的优点是不必对解释变量在预测期内的取值做任何预测。

基于上述VAR模型的优点，同时本文主要是通过计量模型预测误差分析方法来评估经济时间序列数据质量，而不分析经济指标之间的结构关系和参数具体的经济含义等，因此本文选用无约束VAR模型刻画多个经济时间序列数据之间复杂的内在关联规律。

### **3.2.3** 模型预测准确性评价

一般评价一个计量模型拟合的较好，只要拟合模型通过传统的统计和经济意义上的检验即可；但是若要运用拟合模型的预测值进一步做分析应用，还要对拟合模型进行预测准确性评估。所谓模型预测准确性评估就是通过比较动态模型的预测值和样本观测值之间的接近程度，包括微距性和无偏性。

所谓微距性是指动态模型预测值和时间序列样本数据之间的误差不大，在某个标准之内；无偏性是指动态模型预测值和时间序列样本数据所产生的误差序列应该是随机分布在

0值两侧的。

模型预测准确性的评估方法：

1）微距性评估：可用后文离群值评估方法——预测误差百分率评估法对其微距性进行评估；

2）无偏性评估：当模型预测值是连续时，在此运用秩和检验或根据残差白噪声性来对模型预测效果的无偏性进行显著性检验。具体秩和检验方法的实施方法是：假设存在序

列x，若模型预测值大于样本值，则x=1；若模型预测值小于样本值，则x=-1；否则，x=0；然后对序列x进行秩和检验，若显著，则说明模型预测效果无偏；否则，则模型预测有偏。

如果所构建的动态模型的预测准确性评估通过微距性评估和无偏性评估，则表明该模型具有可靠的预测效果；否则，说明该模型预测效果不可靠，要重新构建相应的动态模型。

## 3.3 一致性评估准则

本文从一维和多维两个角度分别对经济指标时间序列建立动态模型，在每个动态模型通过一系列的统计、经济等检验和模型预测效果检验后，在假设样本数据准确的条件下，该动态模型能很好的刻画出该时间序列数据的趋势特征，说明模型预测结果是可靠的。为了进一步运用两个模型预测结果的综合值进行时间序列数据质量诊断，还要对两个动态模型的拟合效果和预测结果进行一致性评估。

百度百科上一致性定义如下：

一致性：英文conformity：校准曲线接近规定特性曲线时的吻合程度。 由上述一致性定义可知，如果两个方法的评估结果是一致性的，则需满足两个条件：

一是两种方法对统计指标相同时期的评估结果差异程度较小，即微距性；二是两种方法对统计指标整个序列的评估结果之差随机分布在0点两侧，即无偏性。根据定义，模型预测效果准确性评估其实就是一致性评估，即校准模型拟合曲线接近样本观测值曲线时的吻合程度。当一维和多维动态模型的预测效果均通过“一致性”评估后，则可以对两个模型的拟合值序列进行一致性评估，这时因不存在校准曲线和规定特性曲线，所以仅进行一致性的微距性检验即可。

一致性评估准则通常有以下几种：

假设一维、多维动态计量模型的第*i*期样本预测值为*Y*1*i*、*Y*2*i*（*i*1, 2,, *n*），其中样本量为*n* 。

### **3.3.1** 距离法

由于从时间序列指标不同的角度所构建的模型对指标*t*期的预测值可能出现不一致，因此，可以构造如下一致性系数：

*t* 

*Y*ˆ

1(*n*1) 2(*n*1)

 *Y*ˆ

(*Y*ˆ

1(*n*1) 2(*n*1)

* *Y*ˆ

) 2

ˆ

其中，*Y*

#### 1 (*n*　1)

ˆ

2(*n*1)

，*Y*

分别为两个计量模型第*n*1期的预测值。

一致性系数*t*描述两种模型对于某时间序列预测结果的差异程度。通过一致性系数可以判断两种模型对相关指标的预测效果是否一致。在给定临界值（如5%）时，

#### 3 动态评估方法的基本思想

若*t* > 5%，则两种模型对该统计指标预测效果的一致性较差，这种一致性评估准则简单直接，但是仅仅对样本外一期的预测值进行一致性检验，不能排除随机性因素；阈值的确定没有公认的标准，不同的阈值可能会得到不同的结论。

### **3.3.2** 变异系数法

由一致性的定义可知，一致性程度就是衡量根据不同角度所构建的模型对相应

经济指标第n+1期预测值*Y*ˆ

1(*n*1)

ˆ

2(*n*1)

和*Y*

的差距，又因为经济指标的预测值都是正的，

故可以用变异系数表示，公式如下：

*C* *S*

2  1 









*Y*ˆ





1( *n* 1) 2 ( *n* 1)

*Y*ˆ







 *Y*ˆ

 

 1( *n* 1) 2 ( *n* 1) 

 *Y*ˆ )  2 

2

 



（3.1）

*v* 1ˆˆ

2 *Y*1( *n*1) *Y*2 ( *n*1) )

其中： *S* 是观测值*Y*ˆ

1(*n*1)

ˆ

2(*n*1)

和*Y*

的标准差。

两观测值的一致性程度越高，式（3.1）的值越小，其充要条件是：

*Y*ˆ*Y*ˆ

*C* 

1(*N*1) 2(*n*1)

*Y*ˆ*Y*ˆ)2

（3.2）

 1(*n*1) 2(*n*1) 

2

越大越好。

简单计算可知，*C*（0,1]，即用变异系数法衡量的一致性程度在0~1之间。*C*值越接近于1越说明一致性程度越高；否则，说明一致性程度越差。一般判别标准是， 当*C*<0.5认为两个预测结果的一致性较差；当0.5*C*0.8认为两个预测结果一致性较好；当*C*0.8认为两个预测结果一致性很好。

这种一致性检验准则是借鉴经济系统中衡量两个系统某时期协调发展程度的协调度系数准则，因此该准则和上述距离法一样，都是仅能衡量某时期的两个不同预测结果之间的一致性程度，且没有进行相应的显著性检验，但是比上述距离法在判别准则的设定上较为合理。

### **3.3.3** 一致性相关系数法

两种计量模型方法对同一样本数据分别进行拟合，为了验证两个模型拟合效果的一致性，构造一致性相关系数*rc* :

西安财经学院硕士学位论文

*Rc* 

2*S*12

*S* 2*S* 2( *y*

*y*) 2

1 2 1 2

其中，*S* 2, *S* 2, *y*, *y*, *S* 分别为两种计量模型样本拟合值方差和均值，计算公式如下：

1 2 1 2 12

1 ˆ

*n*

1 

2 1 

2

*Y*1  *n*

*i*1

*Y*1*i*, *y*2  *n*

*i*1

*n*

*Y*2*i*, *S*1



*n i*1

*n*

(*Y*1*i* *y*1) ，

2 *n*2 1 *n*

1

*S*2  (*Y*2*i**y*2 )

*n*

*i*1

，*S*12  (*Y*1*i* *y*1 )(*Y*2*i* *y*2 ) 。

*i*1

*n*

一致性相关系数*rc* 的取值范围是（-1, 1），一般判别标准是，当*rc*> 0.85 时，说明两种模型的拟合效果一致性很好；当0.5*rc* 0.85 5时，说明两种模型的拟合效果一致性较

好；当*rc* <0.5时，说明两种模型的拟合效果一致性较差。

为了更进一步增加一致性检验的统计效果，降低样本对检验结果的随机影响，还要进行统计显著性检验。假设总体相关系数为，*rc*是其估计值，构造检验统计量来检验总体相关系数=0的原假设是否成立。

统计量构造形式如下：

 *Z*

** 2

*Z*

1 1*r* 2 1 1*r* 2

4*R*3 (1*r* ) *d* 2 2*r* 4 *d* 4

其中：*Z* 

Ln *c*, *Z* [

 *c* c*c* ]

2 1 *r*

*N*2 1*r* 2

*R*(1*r* 2 ) 2

*R* 2 (1*r* 2 )

（*D* 

*c*

*Y*1*y*2, *r* 

*S* 2 *S* 2

1 2

*S* 2 *S* 2

1 2

*S*12 )。

*C* c c 2

在应用中一般先做假设检验，判断两种方法的拟合值是否有显著性一致性相关关系，若显著，则继续用上述的判别标准，判断具体的一致性相关程度；否则，说明这两个方法的检验结果没有显著的一致性相关关系。本文选择一致性相关系数*rc* 对两个模型的拟合效

果进行一致性检验。

由上述各个一致性评估准则的对比分析可知，可用一致性相关系数法对两个动态模型的拟合效果进行一致性评估，用变异系数法对两个动态模型样本外一期预测结果进行一致性评估。在两个一致性评估都通过的情况下，对相应时间序列样本外一期的观测值进行质量评估；否则，表明相应的时间序列样本数据可疑，需要对相应的时间序列样本数据进行审查和处理。

## 3.4 数据质量评估

本文通过时间序列样本数据的趋势特征和复杂的内在关联规律分别建立一维和多维两种动态模型，并对模型进行一系列的检验，在模型通过各种检验、具有良好统计预测功

能的基础上，综合两个模型的预测值，通过综合预测值与实际观测值的差异程度对时间序列数据质量进行评估，找出离群值，并进行异常值显著性检验，最终评估离群值是否可靠。

### **3.4.1** 离群值评估方法

基于计量模型的预测误差离群值评估方法主要是通过模型预测值和实际观测值之间的误差程度来评估该观测值是否离群。具体有以下几种评估方法：

#### 1）相对误差评估法

相对误差是指统计指标第n+1期的实际值与预测值之差与第n+1期实际值的比值。公式表示为：

*e**Yn*1*Y*ˆ1

*n*

*Yn*1

其中，为误差范围阈值（如常用取值1%、5%、10%等），当*e* 时，则初步判定该统计指标的当期实际观测值可靠；否则认为该统计指标第n+1期数据质量异常。这种方法诊断统计数据质量直接简单，但是阈值的选取没有公认的标准，不同的阈值可能会得到不同的结果，甚至完全相反的结果。

#### 2）F统计量评估法

如果经过上述的预测误差叛变表明存在离群观测值，则继续通过统计显著性方法判断该离群观测值是否显著，若显著，则可认为该离群观测值是异常的；否则，认为该离群观测值是可靠的。

分别估计*P*1,... *Pn*的标准差ˆ和*P*1,... *Pn*1的标准差，并计算方差比*F* ：

ˆ  1 (*P**P*) 2, *P*1 *P*

*n*

*n*

*j*

*n*1

*j*

*j*1

*N j*1

*n*1 *n*1

1  1 

ˆ\* (*P**P*) 2, *P* *P*

*N j*1

*j*

*F*() 2

ˆ\*

*N* 1

*j*

*j*1

在显著性水平下，根据*P* {*F**F*}，样本容量为*n*和*F*；可以进一步判断是否是统计意义上的异常值。

#### 3）Z统计量评估法

如动态模型的残差序列满足正态分布，则由正态理论可知动态模型中的因变量也服从正态分布，则可以建立Z统计量。

在显著性水平下，假设模型残差序列的标准差为, *y*ˆ*n*1为统计指标第*n*1期的模型预测值，构造如下区间：

( *y*ˆ*n*1*Z* 2, *y*ˆ*n*1*Z* 2ˆ )

若该统计指标第*n*1期的实际观测值*yn*1落在上述区间内，则说明*yn*1是可靠的；否则，说明*yn*1是异常的。

针对落在置信区间外部的数据还要结合实际情况进行分析，以确定其实际观测值是人为的造假或疏忽，还是真实的数据。

这种方法同样对于显著性水平没有公认的标注，不同的显著性水平诊断的结果可能不同，甚至完全相反。

#### 4）误差百分率评估法

误差百分率评估法核心是计算误差百分率。首先根据动态模型计算时间序列样本内的预测值*Y*ˆ，并计算误差百分率：

*t*

*Y*  *Y*  100%(*t*  1, 2,..., 32) ,

ˆ

*t* t

*t*

*Y*

*t*

由于动态模型是通过了各种统计检验并且具有较好的预测效果的，因此对于离群数据判断的一个基本根据是：通过动态模型计算出的预测值是可靠的。在这种情况下，若全部符合下列条件，假设动态模型预测效果是无偏的，则说明模型能很好的反映该地区经济系统的运行情况，预测结果是可靠的。

A、max {|*t* |}10%

B、1||5%

*t i*

*n*

1*t**n*

*N t*1

在此基础上用该动态模型对时间序列第n+1期观测值进行预测，评估其第n+1期实际观测值的离群性，评估准则如下：

若min {*t*}*n*1 max {*t*}，则认为实际观测值的偏差在历史样本数据的误差百比分

1*t**n* 1*t**n*

率区间内，说明在假设历史样本数据可靠的前提下，该时间序列的第n+1期实际值*Yn*1是可靠的；

若*n*1 max {*t*}或*n*1 min {*t*}，则初步认为*Yn*1是离群值，还需进行数理统计方法

1*t**n* 1*t**n*

统计意义上的显著性水平检验，以进一步判断该离群值是否可靠。

### **3.4.2** 异常值显著性评估准则

常见的传统的异常值显著性评估准则有：3准则、Grubbs准则、Dixon准则、Chauvenet准则、Nari准则和偏度——峰度检验，由文献综述中对异常值评估准则的综述中可知这些准则都有一些明显的弊端，为了弥补这些缺陷，提高异常值评估效果的准确性，文章采用样本分位数检验法[24]对所识别的离群值进行异常值统计显著性检验，和前面传统的方法相比，有计算简便、抵抗异常值污染能力强和可连续使用检验多个异常值等优点。样本分位数检验法具体实施步骤如下：

假设*Y*,, *Y*为服从*N* (*u*,2)的历史样本数据，记*Y*  *Y*

为其顺序统计量，构造

#### 1 *N* (1) (*n* )

异常值评估统计量（包括上侧、下侧和两侧三种情况），当怀疑*Y*（1）（或*Y*（*n*））值异常时，计算上侧（或下侧）异常值评估统计量；否则，当无法确定哪侧有异常值时，计算双侧异常

值评估统计量。当由样本数据计算出的相应统计量的值大于相应的临界值时，在给定的显著性水平（如1%、5%、10%等）下，可判定该离群值为异常值；否则，不能确定该离群值是异常的。该方法所需的各检验统计量如下：

1）当样本标准差已知时，采用0.3样本分位数法，其中：

*Y*1 [*Y*

*Y* ]

上侧检验统计量：*Gn* 

(*n*)

#### 2 (*N*1) (*n*2) ，

#### 1 [*Y*

*Y* ] *Y*

下侧检验统计量：*G*1  2

(*N*1) (*n*2) (1)

，

Max | *Y*

#### 1 [*Y*

*Y* ] |

双侧检验统计量：*MRG**i*1,, *n*

0.3*n*, 0.3*n*为整数

(*I* )

#### 2 (*N*1) (*n*2)。

其中，*n*1  

[0.3*n*] +1,0.3*n*非整数

，*n*2 *n**n*1 1 .

2）当样本标准差未知时，采用0.25样本分位数法，其中：

*Y*1 [*Y*

*Y* ]

上侧检验统计量：*Hn* 

(*N*) 2

*Y*

(*N*3) (*n*4 )

，

 *Y*

(*N*4) (*n*3 )

#### 1 [*Y*

*Y* ] *Y*

下侧检验统计量：*H*  2

(*N*3) (*n*4) (1)

，

#### 1 *Y*　 *Y*

(*N*4) (*n*3 )

Max | *Y*

#### 1 [*Y*

*Y* ] |

双侧检验统计量：*MRH**i*1,, *n*

(*I*) 2 (*n*3) (*n*4)。

*Y*(*n*)*Y*(*n* )

4 3

0.25*n*, 0.25*n*为整数

其中，*n*1  

[0.25*n*] +1,0.25*n*非整数

，*n*4 *n**n*3 1 .

通过比较具体情况下统计量的值和编制好的样本分位数检验法的临界值（临界值表见朱宏（1989）[24]），可判断异常值的统计显著性，进一步评估统计数据是否有异常。

这种方法对于初步诊断出来的离群值进一步进行异常值统计显著性检验，可以有效的

提高统计数据异常值诊断的精度，运用上述样本分位数检验法的前提是假定样本数据序列服从正态分布，因此，在该方法进行统计数据异常值评估时要对相关数据进行正态分布检验。

对统计意义上显著的异常值继续进行该异常值背后的政策变化或自然灾害等突变因

素进行分析，进一步平判断其是现实情况冲击造成的真实数据还是人为的虚假数据。

# 4 动态评估方法的实证分析

文章以1978-2010年间我国某地区的部分主要宏观经济指标作为研究对象，用动态综

合评估方法对2011年该地区相应宏观经济指标数据质量进行评估。

## 4.1 基本情况说明

### **4.1.1** 研究变量

本文的实证分析的主要目的是介绍如何用动态综合评估方法进行宏观经济时间序列数据质量评估，为了使分析过程简单明了，选择最能反映地区经济状况的经济指标——地区生产总值和地区固定资产投资总额作为研究变量，分别记为：zz和tz。

### **4.1.2** 数据说明

文章选择1978-2010 年该地区的地区生产总值和地区固定资产投资总额的观测值作

为历史样本数据，对该地区2011年相应的宏观经济指标数据质量进行评估。为了消除价格因素带来的膨胀误差、减少舍入误差和降低指标序列的波动性对相应指标数据进行转换处理：先转换为以1978年不变价格的地区生产总值和地区固定资产投资总额，再对其做对数变换处理[52]。转换处理后的地区生产总值和地区固定资产投资总额分别记作lnzz1 和

lntz1。数据来源《某地区统计年鉴2012》，具体相关宏观经济指标数据见附录。

### **4.1.3** 基本假设

#### 1）假定所研究宏观经济时间序列历史数据是绝大部分是真实可靠的。

该假设的目的是用历史样本数据来构建动态模型，如果每个从不同角度构建的模型通过各种检验，但是不能通过两个模型预测效果一致性检验，则说明所研究变量的历史数据质量不可靠；否则，继续对该地区2011年相应经济指标数据质量进行评估。数据来源于国家和地方统计局等权威机构，因此“所研究的历史数据绝大部分是可靠的”这一假定是合理的。

#### 2）任何一个宏观经济指标的时间序列都存在着趋势规律，该规律可以通过构建一维动态时间序列模型刻画。

任何一个宏观经济指标在其宏观经济体系中都有其特有的趋势发展规律，其发展的趋势过程就是一个时间序列过程，因此经济指标的发展变化能够通过时间序列模型模拟其发展的趋势规律。

#### 3）所选择的两个宏观经济指标之间存在内在关联关系，该关系可以通过多维动态模型刻画。

由于所选择的两个经济指标来自于同一个宏观经济系统中，同一系统中的各个因素都是相互影响相互制约的，这种复杂关系符合一定客观规律且是动态发展的，因此各宏观经济指标之间的这种内资关联关系能通过多维动态模型给予刻画。

## 4.2 一维动态模型的建立

对该地区的地区生产总值和地区固定投资总额1978-2010年的时间序列数据进行趋势分析，通过ARCH-LM检验可知，两个时间序列数据都没有显著的条件异方差性，因此不选用ARCH模型簇进行趋势分析，通过确定分析、随机性分析和组合分析表明构建确定性趋势组合模型是相对最优的。具体分析如下：

### **4.2.1** 地区固定资产投资总额序列一维动态模型的建立

#### 1）序列趋势分析

由序列{lntz1}的趋势图可知，随着时间推移，序列{lntz1}近似呈上升趋势，可以考虑用确定性趋势回归拟合序列{lntz1}的趋势变化。

用ADF方法对序列{lntz1}进行单位根检验，在5%显著性水平下，通过AIC最小化准则，选择最优滞后阶数，检验结果表4.1所示：

表4.1 序列{lntz1}的ADF单位根检验

Null Hypothesis t-Statistic Prob \*

LNTZ1 has a unit root -0.105108 0.9924

DLNTZ1has a unit root 0.627716 0.8455

D(LNTZ1,2) has a unit root -5.329862 0.0000

由表4.1可知，序列{lntz1}二阶单整，可以考虑用关于时间的二次回归模型或ARIMA

模型进行趋势拟合。

#### 2）建立回归模型ln *tz*1 　*a*　*bt*　*ct* 2 　*u*，其中*t*表示时间趋势项（1978年，*t*　1）。

*t* t

通过Eviews软件拟合模型如下：

Ln *tz*1*t* 3.3233920.004022 *t**ut*

t值74.82044 42.42847

其中，*R*20.983071, *F*1800.175, *DW*0.519038，可知在5%显著性水平下，回归模

型拟合和系数都很显著，但是DW值太小，说明残差可能存在相关性。因此需要对残差进行自相关检验，这里采用Q 统计量进行自相关检验，选择滞后阶数p=8（[32/4] =8），检验结果如图4.1所示：



图4.1 残差Q统计量自相关检验

由图4.1可知，该残差序列是一个近似的AR（1）过程，可以尝试对残差进行AR（1）建模分析。

#### 3）最终模型

首先对回归模型的残差序列进行单位根检验，在5%显著性水平下，通过AIC最小化准则，选择最优滞后阶数，检验结果如下：

表4.2 回归残差的ADF单位根检验

Null Hypothesis t-Statistic Prob \*

LNTZ1 has a unit root -2.662602 0.0094

由表4.2可知，回归残差序列平稳，结合图4.1可知，可以用AR（1）模型进行拟合。结合回归方程和AR模型，对序列{lntz1}构建如下组合模型：

Ln *tz*13.3932840.003941*t*2 *u*

*t*

 *t t*值 32.04971 21.25967

（1978 年， *t*  1）

*Ut* 0.689058*ut*1 *vt*

*t*值5.732962

或者直接写为：

Ln *tz*13.3932840.003941*t*2 0.689058(ln *tz*13.3932840.003941(*t*1) 2) *v*

（1978年，

*t t* 值

*t*1)

32.04971 21.25967 5.732962

*t* 32.04971 21.25967 *t*

其中*R*20.992554, *AIC*1.41, *F*1932.735, *DW*2.066，可知模型拟合和系数均显著、残差不存在一阶自相关性，通过对该模型的残差序列进行Q统计量和ARCH-LM检验，可认定残差序列为白噪声序列且不存在条件异方差性；模型的特征根的倒数（Inverted AR Roots）为：0.69，在单位圆内，因此模型满足稳定性要求；进一步对最终残差序列进

Jarque-Bera正态性检验，得到均值为1.49e-13、JB统计量值为0.498264，对应的P值为

0.779477，因此，在5%显著性水平下，不能拒绝原假设，可认为最终残差序列近似服从均值近似为零的正态分布，说明因变量的信息绝大部分能够被自变量所解释，在统计意义上模型具有良好的预测效果。

#### 4）最终模型预测效果检验

模型预测效果包括两个方面：（1）微距性；（2）无偏性。

##### （1）微距性

由序列{lntz1}的组合模型计算出1979-2010年间的统计指标*Y* (这里指ln *tz*1)的拟合值*Y*ˆ，计算每年的误差百分率：

*t*

*Y*  *Y*  100%(*t*  1, 2,..., 32)

ˆ

*t* t

*t*

*Y*

*t*

由1978-2010年间的误差百分率可得：

1

max {|*t* |}0.9459910% 、

32

|*t* |0.0205055%



1*t*32

32 *t*1

同时满足微距性的两个条件，说明该组合模型的预测值较好的近似于真实值。

##### （2）无偏性

由上文所述，序列{lntz1}的组合模型的残差序列{*v*}（*v* *Y**Y*ˆ ）是白噪声序列，

*t* *t* t

可知残差序列{*vt*}随机的分布在0值两侧，即序列{*vt*}的正负具有随机性，得出序列{*t*}的正负也是随机的，因而，最终模型的预测值是无偏的，不存在预测值偏小或者偏大的趋势。

与此同时，为了进一步找出更好的刻画统计指标序列趋势特征的模型，文章同时用

ARMA模型对序列{lntz1}进行一维动态模型拟合，经过Eviews软件分析，拟合AR（4）模型如下：

*D* ln *tz*1*t* 0.147835*ut*

t值



(6.528)

*Ut* 0.304545*ut*20.307282*ut*4*vt*

t值

1.824

1.806

或者直接写为：

*D* ln *tz*1*t* 0.1478350.304545(*d* ln *tz*1*t*20.147835)0.307282(*d* ln *tz*1*t*40.147835)*vt*

t值(6.528) 1.8241.806

其中，*R*2 0.175, *AIC*1.336 .

比较序列{lntz1}的AR（4）模型和其组合模型的*R*2、*AIC*，可知，AR（4）模型的*R*2较小、*AIC*较大，因此在对序列{lntz1}进行趋势拟合时选择组合模型较好。

上述表明，组合模型能够较真实的刻画出序列{lntz1}，由组合模型预测该地区2011

年的lntz1值为7.734636.

4 动态评估方法的实证分析

### **4.2.2** 地区生产总值序列一维动态模型的建立

#### 1）序列{tz}的组合模型建立

用ADF方法对序列{lntz1}进行单位根检验，在5%显著性水平下，通过AIC最小化准则，选择最优滞后阶数，检验结果如表4.3所示：

表4.3

序列{lnzz1}的ADF单位根检验

|  |  |  |
| --- | --- | --- |
| Null Hypothesis | t-Statistic | Prob \* |
| LNZZ1 has a unit root | 11.41998 | 1.0000 |
| DLNZZ1has a unit root | -2.245342 | 0.1955 |
| D(LNZZ1,2)has a unit root | -6.383187 | 0.0000 |

由表4.3可知，序列{lnzz1}二阶单整，可以用二次回归模型或ARMA模型进行趋势拟

合。

类似序列{lntz1}的建模流程，构造序列{lnzz1}的组合模型如下：

Ln *zz*14.8859720.002709 *t*2 *u*

*t*

*T t*值45.65321 18.90633

（1978年，*t*1）

*Ut* 0.823826 *ut*1*vt*

*t*值12.12269

或者直接写为：

Ln *zz*14.8859720.002709*t*2 + 0.823826(ln *zz*1 - 4.8859720.002709(*t*1) 2) *v*

（1978 ，

*t t* 值

*t*1).

45.65321 18.90633 12.12269

*t*1

45.65321 18.90633 *t*

其中*R*20.997608, *AIC*3.173, *F*6048.181, *DW*1.746424，可知模型拟合和系数

均显著、残差不存在一阶自相关性，通过对该模型的残差序列进行Q统计量和ARCH-LM

检验，可认定残差序列为白噪声序列且不存在条件异方差性；模型的特征根的倒数

（Inverted AR Roots）为：0.82，在单位圆内，因此模型满足稳定性要求；进一步对最终残差序列进Jarque-Bera正态性检验，得到均值为2.07e-13、JB统计量值为1.413975，对应的P值为0.5，因此，在5%显著性水平下，不能拒绝原假设，可认为最终残差序列近似服从均值近似为零的正太分布，说明因变量的信息绝大部分能够被自变量所解释，在统计意义上模型具有良好的预测效果。

#### 2）组合模型预测效果检验

模型预测效果包括两个方面：（1）微距性；（2）无偏性。

##### （1）微距性

由组合模型计算出该地区1979-2010年间的统计指标*Y* (这里指lnzz1)的拟合值*Y*ˆ，计算每年的误差百分率*t* ：

*Y**Y*ˆ

*T* *t* t100%(*t*1, 2,..., 32)

*Y t*

1978-2010年间的误差百分率可得：

1

max {|*t* |}0.01203610% 、

32

|*t* |0.0265585% .



1*t*32

32 *t*1

同时满足微距性的二个条件，说明该模型的预测值较好的近似于真实值。

##### （2）无偏性

由上文所述，最终模型的残差序列{*v*}（*v**Y**Y*ˆ ）是白噪声序列，可知残差序列

*t* *t* t

{ *vt*}随机的分布在0值两侧，即序列{ *vt*}的正负具有随机性，得出序列{*t*}的正负也是随机的，因而，最终模型的预测值是无偏的，不存在预测值偏小或者偏大的趋势。

与此同时，为了进一步找出更好的刻画统计指标序列趋势特征的模型，本文同时用

ARMA模型对序列{lnzz1}进行一维动态模型拟合，经过Eviews软件分析，拟合AR（1）模型如下：

*D* ln *zz*1*t* 0.101950*ut*

t值



(6.216701)

（1978年，*t*1）

*Ut* 0.371151*ut*1*vt*

t值2.107560

或者直接写为：

*D* ln *zz*1*t* 0.1019500.371151(*d* ln *zz*1*t*10.101950)*vt* (1978年，*t*1)

t值(6.216701) 2.107560 (6.216701)

其中，*R*20.13222, *AIC*2.815627, *F*4.441808, *DW*1.8 .

比较序列{lnzz1}的AR模型和回归组合模型的*R*2、*AIC*可知，AR模型的*R*2较小、*AIC*较大，因此在对序列{lnzz1}进行趋势拟合时回归组合模型较好。

上述表明，组合模型能够较真实的刻画序列{lnzz1}，由组合模型预测该地区2011年的lnzz1值为7.843132。

综上分析表明，统计指标序列{lntz1}和{lnzz1}都不具有异方差性，不宜使用ARCH模型拟合其趋势，最终选择用确定性组合模型刻画统计指标序列{lntz1}和{lnzz1}的趋势特征。

## 4.3 多维动态模型的建立

由上文分析可知，统计指标序列{lnzz1}和{lntz1}都是带有趋势的非平稳序列，均是二阶单整，两序列之间若存在协整关系以及上文构建多维动态模型分析中表明，对多变量构建VAR模型有明显的优势，尤其是在预测效果方面，因此，文章在对统计指标序列{lntz1}和{lnzz1}构建多维动态模型时选择VAR模型。

### **4.3.1** **VAR**模型构建

由Eviews软件，根据LR、FPE、AIC、SC和HQ五种评价标准选择VAR模型的最

优阶数，由表4.4可知，这5个评价指标有4个评价指标认为应建立VAR（1）模型，因此，本文对序列{lntz1}和{lnzz1}建立VAR（1）模型。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | 表 4.4 | VAR 模型滞后阶数选择标准 |  | |
| Lag | LogL | LR | FPE AIC | SC | HQ |
| 0 | -4.069166 | NA | 0.005571 0.485533 | 0.583043 | 0.512578 |
| 1 | 76.44918 | 141.7123\* | 1.23e-05\* -5.635934\* | -5.343404\* | -5.554799\* |
| 2 | 77.17912 | 1.167903 | 1.61e-05 -5.374330 | -4.886779 | -5.239104 |
| 3 | 79.25982 | 2.996212 | 1.91e-05 -5.220786 | -4.538215 | -5.031470 |
| 4 | 82.44590 | 4.078175 | 2.11e-05 -5.155672 | -4.278081 | -4.912265 |
| 5 | 87.43695 | 5.489175 | 2.09e-05 -5.227756 | -4.155145 | -4.930259 |
| 6 | 91.57349 | 4.057478 | 2.26e-05 -5.245879 | -3.978248 | -4.894292 |
| 7 | 96.55430 | 3.984655 | 2.42e-05 -5.324344 | -3.861693 | -4.918667 |
| 8 | 99.01757 | 1.576392 | 3.43e-05 -5.201406 | -3.543735 | -4.741638 |

\*indicates lag order selected by the criterion.

在对非平稳序列建立VAR模型时，要先对序列进行协整检验。协整检验一般有两种常用方法：一是基于回归的残差序列的ADF检验，主要是针对单方程的；二是基于回归系数的协整检验，如Johansen协整检验，是一种进行多方程多变量协整检验的好方法。因此，这里采用第二种方法对序列{lnzz1}和{lntz1}进行协整检验。使用Eviews软件进行

Johansen协整检验时，由序列{lnzz1}和{lntz1}的趋势特征，选择“数据中有确定性趋势，协整方程中有截距项”这种形式，由上述选择建立VAR（1）模型，决定协整检验的一阶差分滞后阶数为0。在协整检验的迹统计量值表明，在5%显著性水平下，序列{lnzz1}和{lntz1}存在协整关系。因此，序列{lnzz1}和{lntz1}可以建立VAR（1）模型。

用Eviews软件对VAR（1）进行参数估计，结果如下：

LNZZ1*t* = 0.997965\*LNZZ1*t* 1 + 0.028944\*LNTZ1*t* 1 - 0.020187

 *t*值 8.18952 0.32071 -0.07261

LNTZ1*t* = 0.712378\*LNZZ1*t* 1 + 0.522246\*LNTZ1*t* 1 - 1.693298

 *t*值 2.61184 2.58531 -2.72128

其中，两个单独方程的*R*2值分别为：0.997、0.992，误差平方和分别为：0.0766, 0.3837，

F统计量值分别为：5133.9、1911.8；整体方程的残差协方差行列式为1.97e-05，对数似然估计值为：82.56。由此可知，每个单独的方程和整体方程拟合较好。

对VAR（1）模型进行进行平稳性检验时，其单位根分别为1.04和0.48，有一个单位根落到单位圆外面，说明该VAR（1）模型是不稳定的，但是序列{Lnzz1}和{lntz1}存在协整关系，可以建立VAR（1）模型，并且可以用来做预测。

对VAR（1）模型进行残差检验，一般VAR模型的残差检验包括：残差相关图检验、混合的自相关检验、正态性检验和White异方差检验等。VAR（1）模型残差检验的相应结果如下：由自相关图结果可知，在滞后阶数为8（[32/4] =8）时，VAR（1）模型中两个序列残差之间不存在自相关性，序列{lnzz1}的残差对序列{lntz1}的残差没有交叉相关性，反之， 在滞后1阶和4阶时序列{lntz1}的残差对序列{lnzz1}的残差有轻微的交叉相关性；在滞

后阶数为8的情况下，对VAR（1）模型的残差进行混合的自相关检验时，结果显示在滞后

2阶时，Q统计量值对应的概率最小为0.0514，因此、在5%显著性水平下，VAR（1）模型的残差不存在混合自相关性；用J-B统计量对VAR（1）模型残差进行正态性检验，两个单独方程的残差对应的偏度分别为-0.6878、0.6422，对应的峰度分别为3.1、2.6，对应的J-B统计量值对应的概率分别为0.2816、0.303，并且整体方程的残差的峰度、偏度、J-B统计量对应的概率分比为0.9、0.09、0.295，因此，在5%显著性水平下，该VAR（1）模型的残差均近似服从正态分布；选用含有交叉相的White异方差检验，结果显示整体方程残差、单独方程残差的Chi-sq统计量值对应的最小概率值为0.1454，因此、在5%显著性水平下，

VAR（1）残差不存在异方差性。残差检验结果表明，等号右边变量绝大部分的信息能够被等号左边的变量所解释。

上述表明，所建VAR（1）模型能很好的刻画出序列{lnzz1}和{lntz1}相互之间的内在联系，具有良好的预测效果。

### **4.3.2** **VAR**模型预测效果检验

模型预测效果包括两个方面：（1）微距性；（2）无偏性。

##### （1）微距性

由VAR(1)模型计算出该地区1979-2010年间的统计指标*Y* i （*i*1, 2）的拟合值*Y*ˆi ，

*t* *t*

计算每年相应统计指标的误差百分率*i* ：

*t*

*Y i**Y*ˆ*i*

*i t t*

100%(*t*1, 2,..., 32; *i*1, 2)

*t i*

*Y*

*t*

其中，*Y* 1ln *tz*1, *Y* 2 ln *zz*1 .

*t t t t*

由该地区1978-2010年间相应统计指标的误差百分率可得：

32

|** |  0.020400  5% ；

1

max {|1 |}0.07084910%、 1

*t*

*t*

1*t*32

32 *t*1

max {|2

*t*

32

|}0.02455110%、1 2 |0.0061985% .

|**

*t*

1*t*32

32 *t*1

均同时满足微距性的二个条件，说明VAR（1）模型的预测值较好的近似于真实值。

##### （2）无偏性

由VAR(1)模型的残差检验分析表明，VAR(1)模型的残差序列{*vi* }( *v*i *Y* i *Y*ˆi ，

*t* t t *t*

*i*1, 2）均是白噪声序列，可知残差序列{*vi* }随机的分布在0值两侧，即序列{*vi* }的正负

*t* t

具有随机性，得出序列{*i* }的正负也是随机的，因而，最终模型的预测值是无偏的，不存在预测值偏小或者偏大的趋势。

*t*

综上所述，对序列{lntz1}和{lnzz1}构建VAR（1）能较好的刻画两序列之间内在关系，具有可靠的的预测性。预测该地区2011年lntz1和lnzz1的值分别为7.686854、7.849791。

4 动态评估方法的实证分析

## 4.4 一维和多维动态模型一致性评估

由上文可知，从一维和多维角度分别建立的模型都有可靠的预测性，为了综合两种规律模型的预测结果来对统计数据的质量进行诊断，首先要对两种模型的1978-2010年的拟

合效果进行一致性评估，其次再对两种模型的2011年的预测结果进行一致性评估。如上文所述一致性检验有许多评估方法，本文采用一致性相关系数对两种模型的拟合效果和用变异系数方对两模型的预测结果分别进行一致性评估。

一维和多维动态模型对该地区1979-2010年间指标lntz1和lnzz1的拟合值结果分别如表4.5和表4.6所示：

表4.5 一维动态模型对该地区1979-2010年间的lntz1和lnzz1拟合值

| 年份 | LNTZ1 | LNZZ1 | 年份 | LNTZ1 | LNZZ1 |
| --- | --- | --- | --- | --- | --- |
| 1979 | 3.135405 | 4.484464 | 1995 | 4.560912 | 5.627784 |
| 1980 | 3.175190 | 4.601382 | 1996 | 4.594245 | 5.693996 |
| 1981 | 3.328167 | 4.575952 | 1997 | 4.685924 | 5.782935 |
| 1982 | 3.188835 | 4.619424 | 1998 | 4.818591 | 5.887857 |
| 1983 | 3.371730 | 4.695911 | 1999 | 5.073731 | 5.999217 |
| 1984 | 3.410905 | 4.773226 | 2000 | 5.237499 | 6.117634 |
| 1985 | 3.591933 | 4.909564 | 2001 | 5.438004 | 6.260323 |
| 1986 | 3.821571 | 5.027006 | 2002 | 5.597777 | 6.383898 |
| 1987 | 3.875633 | 5.113990 | 2003 | 5.766763 | 6.517337 |
| 1988 | 4.013764 | 5.192824 | 2004 | 6.018419 | 6.656060 |
| 1989 | 4.033892 | 5.269905 | 2005 | 6.201732 | 6.834085 |
| 1990 | 3.952162 | 5.251094 | 2006 | 6.445875 | 7.040348 |
| 1991 | 4.036527 | 5.353824 | 2007 | 6.698482 | 7.211453 |
| 1992 | 4.164582 | 5.446080 | 2008 | 6.972183 | 7.363481 |
| 1993 | 4.233565 | 5.494059 | 2009 | 7.203948 | 7.539326 |
| 1994 | 4.524795 | 5.622048 | 2010 | 7.495599 | 7.665862 |

表4.6 多维动态模型对该地区1979-2010年间的lntz1和lnzz1拟合值

| 年份 | LNTZ1 | LNZZ1 | 年份 | LNTZ1 | LNZZ1 |
| --- | --- | --- | --- | --- | --- |
| 1979 | 3.011520 | 4.453400 | 1995 | 4.566609 | 5.626758 |
| 1980 | 3.130775 | 4.589179 | 1996 | 4.594199 | 5.680691 |
| 1981 | 3.210044 | 4.556015 | 1997 | 4.682980 | 5.763348 |
| 1982 | 3.124666 | 4.592780 | 1998 | 4.813963 | 5.865828 |
| 1983 | 3.309375 | 4.681815 | 1999 | 5.040655 | 5.979993 |
| 1984 | 3.383190 | 4.764556 | 2000 | 5.201513 | 6.097608 |
| 1985 | 3.612871 | 4.923496 | 2001 | 5.408521 | 6.244910 |
| 1986 | 3.860372 | 5.060328 | 2002 | 5.565446 | 6.366088 |
| 1987 | 3.945781 | 5.151629 | 2003 | 5.735200 | 6.498341 |
| 1988 | 4.085177 | 5.235330 | 2004 | 5.969498 | 6.639209 |
| 1989 | 4.130938 | 5.310691 | 2005 | 6.183299 | 6.823557 |
| 1990 | 4.013897 | 5.264352 | 2006 | 6.464939 | 7.043408 |
| 1991 | 4.125157 | 5.370964 | 2007 | 6.719910 | 7.219765 |
| 1992 | 4.257791 | 5.465463 | 2008 | 6.971688 | 7.372639 |
| 1993 | 4.304683 | 5.502586 | 2009 | 7.209595 | 7.551343 |
| 1994 | 4.586521 | 5.644706 | 2010 | 7.447569 | 7.671573 |

由上文可知，选择一致性相关系数*rc* :

*Rc* 

2*S*12

*S* 2*S* 2( *y*

*y*) 2

1 2 1 2

其中，*S* 2, *S* 2, *y*, *y*, *S*分别为两种计量模型样本预测值的方差、均值和协方差，，计算公式如下：

1 2 1 2 12

*n*

1 ˆ

*n*

1 

2 1 

2

*Y*1  *n*

*i*1

*Y*1*i*, *y*2  *n*

*i*1

*Y*2*i*, *S*1



*n i*1

*n*

(*Y*1*i* *y*1) ，

2 *n*2 1 *n*

1

*S*2  (*Y*2*i**y*2 )

*n*

*i*1

，*S*12  (*Y*1*i* *y*1 )(*Y*2*i* *y*2 ) 。

*i*1

*n*

先对两组序列{lntz1}的拟合值进行一致性评估，不防把一维模型和多维模型对序列

{lntz1}的拟合值分别看作一致相关系数*r*中的*Y* 1、*Y* 2 ( *t*1, 2..., 32 )，经计算得，

*c* t t

*Y*1 4.7708856411 ,

*Y*2 4.7708856875 ,

*S* 21.5805391165

， *S* 2  1.5810182515 ，

*S*12 1.5810182515, r*c* =0.9994203229 .

1

2

根据所求得两组序列{lntz1}拟合值的一致性相关系数为0.999420323> 0.85, 根据上文所述的判别准则，可初步判断这两种模型的拟合效果一致性非常好。

为了更进一步增加一致性评估的统计效果，降低样本对检验结果的随机性影响，还要进行统计显著性检验。假设总体相关系数为，*rc*是其估计值，构造检验统计量来检验总体相关系数=0的原假设是否成立。

统计量构造形式如下：

 *Z*

** 2

*Z*

1 1*r* 2 1 1*r* 2

4*R*3 (1*r* ) *d* 2 2*r* 4 *d* 4

其中：*Z* 

Ln *c*, *Z* [

 *c* c *c* ]

2 1 *r*

*N*2 1*r* 2

*R*(1*r* 2 ) 2

*R* 2 (1*r* 2 )

（*D* 

*c*

*Y*1*y*2, *r* 

*S* 2 *S* 2

1 2

*S* 2 *S* 2

1 2

*S*12 )。

*C* c c 2

计算得，*d*-8.5278E-9、r=0.9997305611、z=4.4145047886、20.0306723887 、

*z*

*u*25.206247288，*u*统计量值对应的概率P值小于0.01，因此，在5%显著性水平下，拒绝原假设，可以认为两个序列是强相关的，又因为两组lntz1拟合值的一致性相关系数为0.9994203229> 0.85, 可认为两种模型的拟合效果一致性非常好。

同理，可检验序列{lnzz1}的两组拟合值的一致性，相关计算结果分别为：*y*1 =

5.7816359062 、

*S*

*S*

2

2

*Y*2 =5.781635824、1

=9.4451730139、2

=9.5564728657 、

*S*12

=9.4992306086、*rc*

= 0.999832400、*d* =8.657293E-9、r =0.9998495527、z =

4.6934989958、2 =0.0299212494、*u* =27.133565564, 可知*r* = 0.999832400> 0.85, *u*

*z* c

统计量值对应的概率P值小于0.01，因此，在5%显著性水平下，可认为两组模型对序列

{lnzz1}的拟合效果高度一致性。

一维动态模型对该地区2011年统计指标lntz1和lnzz1的预测值分别为7.734636、

7.843132，多维动态模型相应的预测值分别为7.686854、7.849791。其两种2011年统计指标lntz1和lnzz1预测值的变异系数法一致相关检验结果分别为：*C*1=0.99998, *C*2=0.9999；由上文变异系数法的评判标准可知，两个模型2011年相应统计指标的预测结果一致性非常好。

综上表明，在假设历史数据可靠的情况下，所建立的一维动态模型和多维动态模型都能刻画出序列{lnzz1}和{lntz1}的趋势特征和内在关联规律，具有可靠的预测性，且拟合效果和预测结果高度一致，因此，为了进一步提高对序列{lnzz1}和{lntz1}的预测精度，可以综合上述两个模型的预测结果，如两者的算术平均值等，来提高运用模型预测误差分析方法评估数据质量的效果。

## 4.5 数据质量评估

由上文可知，两个动态模型通过各种检验、拟合效果和预测结果的一致性评估，说明可以运用动态综合方法对2011年该地区的地区生产总值和地区固定资产投资总额观测值进行质量评估。相关评估分析见表4.7：

表 4.7 2011年该地区相关经济指标数据质量评估分析表

| 经济  指标 | 一维动态模  型预测值 | 多维动态模  型预测值 | 综合预测值 | 实际观测  转换值 | 误差百分比 | 数据质量可靠  性评估 |
| --- | --- | --- | --- | --- | --- | --- |
| Lntz1 | 7.734636 | 7.686854 | 7.710745 | 7.636602 | -0.00970881 | 可靠 |
| Lnzz1 | 7.843132 | 7.849791 | 7.8464615 | 7.858379 | 0.0015165985 | 可靠 |

说明，表4.7中综合预测值是一维动态模型预测值和多维动态模型预测值的算术平均值，由上文的分析表明，其算术平均值可以用来作为该地区2011年统计指标lntz1和lnzz1的预测值且是可靠的；实际观测转换值是官方公布的相应的统计指标观测值，换算成1978年不变价格的自然对数数值。

由上文可知一维和多维动态模型对应的max {|1 |}分别为0.94599、0.070849，由表

1*t*32 *t*

## 4.7 可知，lntz1 的误差百分比绝对值为|1

| =0.00970881，介于每个模型的误差百分比率1

之中；一维和多维动态模型对应的max {|2 |}分别为0.012036、0.024551，lnzz1的误差

*t*

1*t*32 *t*

百分比绝对值为|2

| =0.0015165985，介于每个模型的误差百分比率2 之中。根据前文所

述统计数据异常性评价准则，在历史数据假定可靠的情况下，认为公布的该地区2011 年

*t*

西安财经学院硕士学位论文

的地区生产总值数据和地区固定资产投资总额数据是准确的。由于2011年该地区相关统计指标数据质量评估结果是可靠的，因此不需要进一步进行样本分位数法进行异常值显著性检验。

# 5 结论与展望

## 5.1 结论

由于描述宏观经济系统的指标很多，加上有些经济指标的数据不易得到，文章仅选取比较有代表性的和关联性较强的地区生产总值和地区固定投资总额进行动态综合模型方法评估宏观经济时间序列数据质量的实证分析。在对该地区1978-2010年间的上述每个统计指标进行ARCH-LM检验时发现两个时间序列均没有异方差性，故不采用ARCH模型进行趋势拟合，对比确定性趋势组合模型和ARIMA模型的拟合效果，结果表明确定性趋势组合模型能更好的刻画这两个时间序列的趋势特征。

通过对确定性趋势组合模型和VAR模型进行一致性检验，结果表明两个模型的预测效果高度一致，进而综合两个模型的预测值，运用综合预测值误差分析对2011年该地区的地区生产总值和地区固定投资总额的准确性进行评估，结果表明相关的统计数据质量是可靠的。

动态评估方法主要适用于对多个宏观经济时间序列数据的质量进行评估。分析单个宏观经济时间序列的趋势特征，构造一维动态模型拟合其趋势；分析多个宏观经济时间序列之间的关联规律，构建多维动态模型刻画其内在机制，综合不同模型的预测值，该综合预测值能平衡一维和多维动态模型的预测误差，其综合的预测效果在理论上比一维或多维动态模型的预测效果较好，通过比较动态预测值和经济指标观测值之间的差异程度，用检验异常值的方法进行统计数据质量评估时更准确，相对单一计量模型方法往往能提高对统计数据质量评估的效果。

在一维和多维动态模型预测效果的一致性检验中和检测异常数据方面，传统的一致性检验方法和异常值检验准则，一般都是事先给出所构造的统计量的一系列取值区间来划分一致性的等级，划分标准不一，这样往往造成没有一个统一的标准，本文采用的一致性相关系数*rc* 进行一致性相关检验、用统计中检验异常值的方法和样本分位数对误差进行统计

上的显著性检验该，能有效解决上述问题，提高相应的检验效果。

本文所述的时间序列数据质量动态评估方法，在理论上和实际应用中还有许多不足：

#### 1）在理论上计量模型往往假定时间序列历史数据是准确的，这点在实际中往往是不

能满足的；

#### 2）部分时间序列数据很难从多个角度进行研究或动态模型的预测效果不佳，这时使得动态评估方法的实施较困难甚至失效；

3）在两个动态模型预测效果一致性检验准则和异常值检验准则的设计上都要求时间序列残差序列服从正态分布，这对于评估非正态分布的时间序列残差数据的质量效果较差，在这种情况下，根据实际情况，可以选择对所评估的统计指标数据进行相应的数据转换或者构建更合适的一致性检验准则和异常值检验准则，以提高评估效果；

4）在两个动态模型的一致性评估不通过时，只能诊断出相应时间序列的历史数据中有异常值现象，但是不能诊断出异常值的具体位置。

## 5.2 展望

鉴于本文所构建的时间序列数据动态评估方法的思想，对非时间序列数据也可以尝试综合多个角度的分析结果对其进行数据质量评估，构建实用性更一般的评估准则增加动态评估对非时间序列数据质量的评估精度。

希望尽快实施国家统计局局长马建堂在中国统计学会第八届一次常务理事会上的讲话中的三点指出：要深入落实科学发展观，紧紧围绕提高统计能力，提高数据质量和公信力，一是要着力强化对提高统计能力的研究，二是着力强化对提高统计数据质量的研究，三是着力强化对提高统计公信力的研究。同时，结合国际数据质量的标准，加强我们宏观经济核算体系与国际接轨，使地方和中央的统计数据核算口径、方法有统一的标准；健全我国的统计机制，完善统计法规，为我国统计工作营造良好的环境提供法律保障；加强我国统计工作人员的业务和素质培训，重视基层统计工作，确保数据统计的质量；加大全国范围内统计的重要性和实用性，营造一个良好的统计文化氛围；加强宏观和微观层次的管理调控和衔接，确保整体经济运行的协调性，协助国家统计部门，保证统计数据的一致性。尽快提升我国统计数据质量的可靠性，为国家和社会谋福利。

参考文献

[1] Perkins D. Refoming China's Economic System[J]. China Economic Review. 1988,26(2).

[2] Rawski T G. What's Happening to China's GDP Statistics[J]. ChinaEconomicReview,2001,12(4).

[3] Rawski T G. China's GDP Statistics A Case of Caveat Lectoi[J]. ChinaEconomicQuarterly,2001,12(5).

[4]聂富强，崔名铠，向荣美。政府统计数据质量内涵的演化与启示[J]. 统计研

究,2011,28(5).

[5] Hansen, M. H., W. N. Hurwitz. The Problem of one response in Sample Surveys [J]. Journal of the American Statistical Association,1946, (41).

[6] Hansen, M. H., W. N. Hurwitz, Bershad, M. Measurement Errors in Censuses and Surveys

[J]. Bulletin of the ISI,1961, (38).

[7] Warner, S. L. Randomized Response A Survey Technique for Eliminating Evasive Answer Bias[J]. Journal of the American Statistical As-sociation,1965, (60).

[8] United Nations. Handbook of Statistical Organization[Z]. Newyork, 1980.

[9] Dalenius. Errors and Other Limitations of Survey, Statistical Methods and the Improvement of Data Quality [M]. London: Academic Press Inc, 1983.

[10] Brackstone, G. Managing Data Quality at Statistics Agency [J]. Survey Methodology, 2000, (25).

[11] Frschl, K. A., Grossmann W. The Role of Metadata in Using Administrative Sources [J].

Research in Official Statistics,2000, (3).

[12] Klein, L. R., and S. Ozmucur. The Estimation of China's Economic Growth Rate[J]. Journal of Economic and Social Measurement, 2002,28( 4).

[13]阙里，钟笑寒. 中国地区GDP 增长统计的真实性检验[J].数量经济技术经济研

究,2005, (4).

[14]刘孝新，朱慧明，胡先红.事后预测及反常结构判断法在统计数据质量评价中的应用

[J].统计研究，1997, (2).

[15]周潮.基于SARIMA模型涉农贷款专项统计数据质量评估[J].甘肃金融,2008，（7）。

[16]杨海ft，许启发. 统计数据质量的逻辑评估方法研究[J].上海统计,2001，（7）。

[17]刘洪，黄燕.我国统计数据质量的评估方法研究——趋势模拟评估法及其应用[J].统

计研究,2007,24(8).

[18]刘洪，黄燕.基于经典计量模型的统计数据质量评估方法[J]. 统计研究，2009，（3）.

[19]张维群，耿宏强.区域宏观经济统计数据质量定量诊断方法及应用研究[J]. 统计与信息论坛,2010，（9）。

[20]赵进文.异常值对计量建模影响的典型案例[J].统计研究,2010，（12）。

[21]卢二坡，黄炳艺.基于稳健MM 估计的统计数据可靠性评估方法[J].统计研究，

2010,27(12).

[22]卢二坡，张焕明.基于稳健主成分回归的统计数据质量评估方法[J].统计研究，

2011,28(8).

[23]楼润瑜，吴江云，王水生，钟继.检验数据异常值判断方法的优选[J].检验检疫科学,2008，（6）。

[24]朱宏. 用样本分位数方法同时检验正态样本多个异常值[J]. 数理统计与应用概率,1989，（4）。

[25]赵慧，甘仲惟，肖明.多变量统计数据中异常值检验方法的讨论[J].华中师范大学学报：自然科学版,2003, 37（2）。

[26]成邦文，董丽娅，杨峻.研究与开发机构统计数据质量与异常点的对数正态分布检验与识别[J].统计研究,2000，（1）。

[27]成邦文，师汉民，王齐庄.多维统计数据质量检验与异常点识别的模型与方法[J]. 数学的实践与认识,2003, 33（4）。

[28]张德然.统计资料中异常值的检验方法[J].统计研究,2003，（5）。

[29]李子奈，周建.宏观经济统计数据结构变化分析及其对中国的实证[J].经济研究，

2005, (1).

[30]王斌会，陈一非.基于稳健马氏距离的多元异常值检测[J].统计与决策，2005，（3）.

[31]朱文杰.多维经济数据质量诊断的“投影寻踪”法[J].统计与决策,2007，（20）。

[32]孙颖等.现今统计制度存在的问题及改革建议[J].金融研究,1996，（12）。

[33]傅德印.浅论政府统计数据质量控制技术体系[J].统计研究,1988，（2）。

[34]傅德印，刘晓梅.贯彻国际标准建立健全统计数据质量管理与保证体系[J]. 统计与信息论坛,2000，（2）。

[35]傅德印，陶然.对政府统计数据质量成本的探讨[J].统计研究,2007，（8）。

[36]孟连，王小鲁.对中国经济增长统计资料可信度的估计[J]. 经济研究，2000，（10）.

[37]傅德印，陶然.对政府统计数据质量成本的探讨[J].统计研究,2007，（8）。

[38]许永洪，曾五一.基于中国城镇统计资料的CPI偏差估计[J].统计研究,2009，（4）。

[39]杨青.统计数据质量研究新思路——误差研究[J].统计研究,2000，（8）。

[40]成邦文，董丽娅.社会经济统计资料的误差分析[J].统计研究,2002，（11）。

[41]王华，金勇进. 统计数据准确性评估的误差效应分析方法[J]. 统计信息 论

坛,2009,24(9).

[42]王珊珊，蔡永生.基于Bayes统计决策的误差测定方法[J].统计与决策,2008，（4）。

[43]归庆明，李涛，衡广辉.时间序列异常值探测的Bayes方法及其在电离层VTEC数据处理中的应用[J].武汉大学学报：信息科学版,2011, 36（7）。

[44]杨青云，赵培英，杨冬青，唐世渭，童云海.数据质量评估方法研究[J].计算机工程与应用,2004，（9）。

[45]朱秀君.政府统计中造假与治假的博弈分析[J].数量经济技术经济研究,2000，（7）。

[46]卢冶飞.政府统计中造假与治假的博弈分析[J].统计究,2003，（5）。

[47]林勇，董梅.统计数据质量问题的成因：基于博弈的分析[J].统计与决策,2008，（4）。

[48]林勇.统计数据博弈：模型、含义与管理对策[J].统计与决策,2006，（17）。

[49]刘海清，熊祖辕.统计信用与统计数据质量研究[J].统计研究,2009，（12）。

[50]李丹玲，陈平雁，周风麒.基于线性*v*支持向量回归机的异常数据检验[J].数理统计与管理,2011, 30（1）。

[51] 高铁梅. 计量经济分析方法与建模—Eviews应用及实例(第二版)[M]. 北京: 清华大学出版社, 2009: 267-268.

[52]王燕.应用实践序列分析（第二版）[M].北京：中国人民大学出版社，2008: 182-183.

# 攻读硕士学位期间发表的论文

在硕士研究生期间取得成果情况如下：

赵晓洋，王杨. 《我国主要宏观经济统计指标数据的协调性探讨》. 中国—东盟博览.2012.（8）.

王杨，赵晓洋. 《我国汽车销量的主要影响因素分析和展望》. 中国—东盟博览.2013.（3）.

致 谢

感谢我的导师张维群教授的关心、指导和教诲。张维群教授追求真理、献身科学、严以律已、宽以待人的崇高品质对我将是永远的鞭策。

作者在攻读硕士学位期间的工作自始至终都是在张维群教授全面、具体的指导下进行的。张维群老师渊博的学识、敏锐的思维、民主而严谨的作风，使我受益匪浅，终生难忘。

同时感谢在研究生三年期间为我授课的所有教师，感谢他们对我的谆谆教诲与指导、关怀与照顾，他们的学术智慧和高尚的人格，是我毕生学习的榜样和奋斗的方向；感谢同学和舍友在生活和学习上对我的关心和照顾。

最后，谨向百忙之中参于评审和答辩的各位老师致以衷心的感谢和敬意！

附录

|  |  |  |  |
| --- | --- | --- | --- |
|  | 附表 1 | 1978-2010 年某地区相关经济指标数据 |  |
| 年份 | 地区生产总值（亿元） | 地区固定资产投资总额（亿元） | 固定资产投资价格指数 |
| 1978 | 81.07 | 20.3548 | 1 |
| 1979 | 94.25 | 21.622 | 1.016 |
| 1980 | 94.91 | 27.802 | 1.064 |
| 1981 | 102.09 | 22.9244 | 1.096 |
| 1982 | 111.95 | 29.4805 | 1.107 |
| 1983 | 123.39 | 30.8266 | 1.124 |
| 1984 | 149.35 | 40.3864 | 1.168 |
| 1985 | 180.87 | 57.9881 | 1.244 |
| 1986 | 208.31 | 63.53 | 1.309 |
| 1987 | 244.96 | 80.8912 | 1.422 |
| 1988 | 314.48 | 94.721 | 1.692 |
| 1989 | 358.37 | 95.1793 | 2.01 |
| 1990 | 404.3 | 103.7154 | 2.042 |
| 1991 | 468.37 | 124.9323 | 2.16 |
| 1992 | 531.63 | 142.4653 | 2.365 |
| 1993 | 678.2 | 228.2062 | 2.644 |
| 1994 | 839.03 | 283.2912 | 3.329 |
| 1995 | 1036.85 | 324.325 | 3.895 |
| 1996 | 1215.384 | 371.9955 | 4.21 |
| 1997 | 1363.69 | 424.0987 | 4.277 |
| 1998 | 1458.4 | 544.8916 | 4.114 |
| 1999 | 1592.64 | 619.2743 | 4.011 |
| 2000 | 1804 | 745.85 | 3.943 |
| 2001 | 2010.62 | 850.66 | 3.908 |
| 2002 | 2253.39 | 974.63 | 3.853 |
| 2003 | 2587.72 | 1278.72 | 3.872 |
| 2004 | 3175.58 | 1544.19 | 3.969 |
| 2005 | 3933.72 | 1982.04 | 3.973 |
| 2006 | 4743.61 | 2610.22 | 4.045 |
| 2007 | 5757.29 | 3642.13 | 4.247 |
| 2008 | 7314.58 | 4851.41 | 4.54 |
| 2009 | 8169.8 | 6563.39 | 4.535 |
| 2010 | 10123.48 | 8561.24 | 4.731 |