



中国大学生计算机设计大赛大数据主题赛

# 绿水青山 富裕乡关

基于DLF-LSTM的乡村发展报告



杨熙承 王雨欣 徐彬斌 唐胡煜 任俊濠

和鲸社区



## 摘要

在新时代背景下，乡村发展不仅是实现乡村振兴战略的关键，更是推动社会和谐与共同富裕的重要途径。本文从多个维度探究乡村发展的多方面内容，包括粮食安全、科技影响、乡村建设、产业结构、乡村就业和人口等方面，旨在提出全面而具体的发展策略。

首先，粮食安全是国家的生命线，保障粮食安全是国家战略的基石。本文通过对近年来的粮食播种面积和主要粮食产量的时间序列分析，突显了我国在确保粮食安全方面取得的显著成就。

在科技影响方面，本文从农业生产和农民生活质量两个层面进行探讨。利用 Holt 指数平滑预测法，展示了科技创新如何为农业生产注入活力，并显著提升农民的生活水平，这也与我国发布的数字农业发展报告相呼应。

乡村建设部分，本文利用复合年增长率（CAGR）深入分析了各省基础设施建设的进展，并提出了针对性的建议。同时，通过量化数据展示了生态环境保护和改善对乡村环境的积极影响。这些努力最终促进了乡村旅游业的发展，与中央农村工作会议提出的美丽乡村建设目标相契合。

在探讨农民生活水平时，本文从农村就业和产业结构变化出发，使用循环神经网络预测了乡村发展的潜在产业方向。通过斯皮尔曼分析收入支出信息并且进行 PCA 和牛顿多元回归、分析恩格尔系数以及基本生活工具，深入了解了乡村居民的生活质量。

最后，在城乡人口方面，本文通过 DLF-LSTM 预测分析指出，随着中国总人口趋于饱和，城市人口稳定，现在正是推动乡村发展、促进农村人口回流的大好时机。国家领导人在推动乡村振兴中的战略思考中，也强调了加快乡村建设，实现城乡融合发展的的重要性。

**关键词：**DLF-LSTM 斯皮尔曼相关性 牛顿多元回归 PCA CAGR Holt



# 目录

<b>1 数据集使用情况</b>	<b>1</b>
<b>2 模型介绍</b>	<b>2</b>
2.1 指数平滑模型 (Holt) . . . . .	2
2.2 LSTM . . . . .	2
2.3 DLF . . . . .	4
2.4 DLF-LSTM . . . . .	4
2.5 模型对比 . . . . .	6
<b>3 背景</b>	<b>7</b>
3.1 两会中的乡村发展 . . . . .	7
3.2 角度分析 . . . . .	8
<b>4 粮食安全问题</b>	<b>9</b>
4.1 农作物播种面积 . . . . .	9
4.2 主要粮食产量 . . . . .	10
<b>5 科技进步</b>	<b>11</b>
5.1 科技对乡村农业的发展的影响 . . . . .	11
5.1.1 科技主要贡献 . . . . .	11
5.1.2 数字农业渗透率 . . . . .	12
5.1.3 农业机械总动力 . . . . .	14
5.2 科技对乡村居民生活的影响 . . . . .	15
<b>6 乡村建设</b>	<b>17</b>
6.1 基础设施建设 . . . . .	17
6.2 基础环境建设 . . . . .	18
6.3 生态旅游建设 . . . . .	18
<b>7 农民生活水平</b>	<b>20</b>
7.1 农村就业 . . . . .	20
7.1.1 就业结构改善 . . . . .	20
7.1.2 就业人员预测 . . . . .	21

7.2 可支配收入 . . . . .	23
7.3 基本生活 . . . . .	25
7.4 幸福指数 . . . . .	26
<b>8 城乡人口结构</b>	<b>28</b>
8.1 城乡人口预测 . . . . .	29
<b>9 结论和展望</b>	<b>32</b>
<b>参考文献</b>	<b>33</b>



# Chapter 1

## 数据集使用情况

表 1.1 已使用竞赛提供数据集

数据集名称	数据来源
主要农作物播种面积和产品产量数据.csv	竞赛官方数据
乡村和城镇人口数据.csv	竞赛官方数据
全国农业主要科技贡献相关指标.csv	竞赛官方数据
农村居民平均每百户年末主要耐用消费品拥有量数据.csv	竞赛官方数据
城乡就业人员数据.csv	竞赛官方数据
《乡村振兴战略规划（2018 — 2022 年）》全文.txt	竞赛官方数据
《全国乡村产业发展规划（2020-2025 年）》全文.txt	竞赛官方数据
农村居民人均收支数据.csv	竞赛官方数据
世界各国乡村人口数据.csv	竞赛官方数据

表 1.2 拓展补充数据集

数据集名称	数据来源
城乡互联网.csv	中国数字乡村发展报告（2020）
农民收入.csv	国家统计局
全国恩格尔系数.csv	数据基地
十四五计划.csv	十四五数字经济发展规划
数字农业渗透率.csv	中国信通网、前瞻产业研究院
乡村环境.csv	城乡建设统计年鉴、Wind、民生证券研究院
乡村基础设施.csv	和鲸平台资源数据集
乡村旅游.csv	《中国乡村旅游发展报告绿皮书》
中国机械总动力.csv	国家统计局



## Chapter 2

### 模型介绍

#### 2.1. 指数平滑模型 (Holt)

Holt 线性趋势方法，也称为 Holt 双参数指数平滑法，是一种时间序列预测技术，用于处理显示出明显线性趋势的数据。Holt 线性趋势方法包含两个方程：一个是水平方程，另一个是趋势方程。

1. 水平方程计算当前时间点的平滑水平值：

$$\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \quad (2.1)$$

2. 趋势方程计算当前时间点的趋势值：

$$b_t = \beta(\ell_t - \ell_{t-1}) + (1 - \beta)b_{t-1} \quad (2.2)$$

其中： $\ell_t$  是在时间  $t$  的水平成分。 $\alpha$  是水平平滑参数， $0 < \alpha < 1$ 。 $y_t$  是在时间  $t$  的实际观测值。 $b_{t-1}$  是在时间  $t-1$  的趋势成分。 $b_t$  是在时间  $t$  的趋势成分。 $\beta$  是趋势平滑参数， $0 < \beta < 1$ 。

最终的预测方程结合了水平和趋势成分来生成对未来时间点的预测：

$$\hat{y}_{t+h|t} = \ell_t + hb_t \quad (2.3)$$

这里， $\hat{y}_{t+h|t}$  是在时间  $t$  基础上对  $h$  期后的预测值。

Holt 线性趋势方法特别适用于那些具有线性趋势、数据量较少但无季节性的时间序列。通过选择适当的  $\alpha$  和  $\beta$  值，可以调整模型对水平和趋势的敏感程度，进而影响预测的准确度。在本报告中，将 2014、2016、2018 年的城乡互联网使用程度的数据通过指数平滑法进行 2020 和 2022 的预测。

#### 2.2. LSTM

LSTM 从被设计之初就被用于解决一般递归神经网络中普遍存在的长期依赖问题，使用 LSTM 可以有效的传递和表达长时间序列中的信息并且不会导致长时间前的有用信息被忽略（遗忘）。

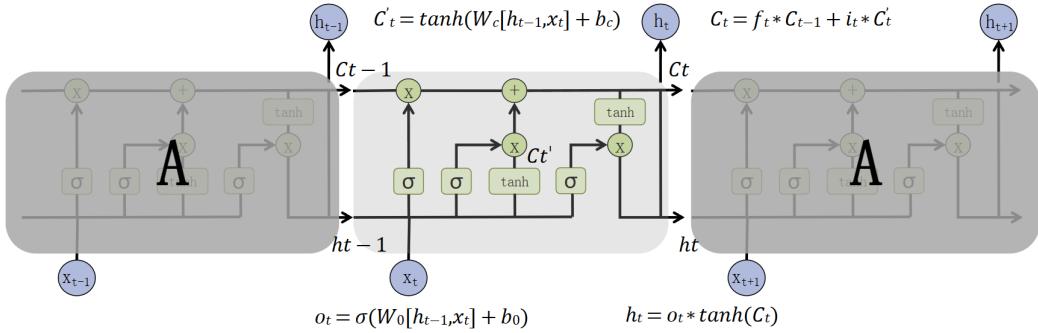


图 2.1 LSTM 原理图

一个 LSTM 单元主要由以下几个部分组成：遗忘门、输入门、细胞更新和输出门。简化图如下：

假设在时间步  $t$ , LSTM 接收到输入向量  $x_t$  和上一个时间步的输出  $h_{t-1}$  以及细胞状态  $c_{t-1}$ 。

1. 遗忘门:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.4)$$

$f_t$  是遗忘门的输出,  $W_f$  是遗忘门的权重矩阵,  $b_f$  是偏置项,  $\sigma$  是 sigmoid 函数, 它确保输出值在 0 和 1 之间, 决定了要从细胞状态中遗忘多少信息。

2. 输入门和候选细胞状态:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (2.5)$$

输入门决定了我们将在细胞状态中更新多少新信息, 而候选细胞状态则提供了一个新的候选信息集, 这些信息可能会被添加到细胞状态中。

3. 细胞状态更新:

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \quad (2.6)$$

当前细胞状态  $c_t$  是前一状态  $c_{t-1}$  经过遗忘门“遗忘”一部分后, 与新的候选细胞状态经过输入门“筛选”一部分后相加得到的。

4. 输出门和输出值:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) h_t = o_t * \tanh(c_t) \quad (2.7)$$

输出门决定了哪些细胞状态的信息将输出, 而输出值  $h_t$  则是这部分信息经过处理后用于传递到下一个 LSTM 单元或用于当前单元的输出。

这些公式共同定义了 LSTM 单元如何通过时间传递和更新信息, 使其能够捕捉长期依赖关系。在本篇报告中, 在预测数字农业渗透率、城乡就业人员、城乡总人口时采用了 LSTM 对接下来几年进行了时间预测。

## 2.3. DLF

双逻辑函数（Double Logistic Function, DLF）是一种适用于受到多重因素影响，既有增长又有减少的进行动态复杂变化的数学模型，比如人口。双逻辑函数通过结合两个逻辑函数，更加灵活地模拟出人口先增长后减少或反之的变化情况，如下图所示：

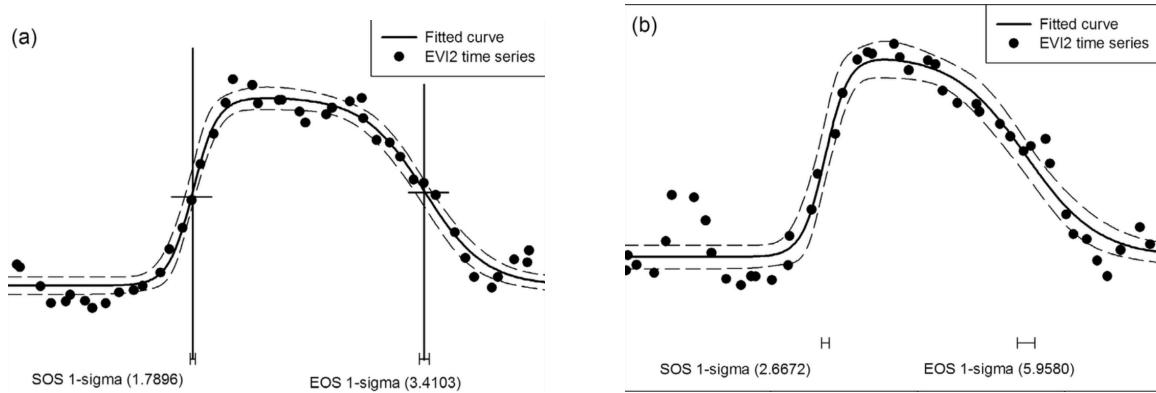


图 2.2 双逻辑函数进行曲线拟合

双逻辑函数的模型公式为：

$$P(t) = \frac{K}{1 + e^{-r(t-t_0)}} + \frac{L}{1 + e^{-s(t-t_1)}} \quad (2.8)$$

其中  $K$  和  $L$  表示增长和减少的幅度,  $r$  和  $s$  是相应的增长率,  $t_0$  和  $t_1$  是拐点时间。

相比单一逻辑函数，双逻辑函数具有灵活性强、精度高和直观性好的优势。假设一个城市在某段时间内经历了快速人口增长，随后由于某些原因（如经济衰退、自然灾害）导致人口减少。利用双逻辑函数，可以分别设定增长和减少的幅度、速度和时间点，精确模拟出这一变化过程，提供有效的预测和分析工具。

## 2.4. DLF-LSTM

由于我们所获得的数据大部分与年份相关，且数据量较少，容易产生过拟合的情况。单纯依赖数学模型难以准确捕捉数据中的依赖关系，也无法拟合随着外部因素动态变化的信息。因此，我们考虑将两种模型进行融合，结合数学模型的精确性和机器学习模型的适应性，以提高整体模型的表现。

在本文中，我们选择双逻辑函数（DLF）和长短期记忆网络（LSTM）作为基学习器，同时处理相同的训练集，将得到的参数放入同一个数据集，并作为输入值输入到元学习器中。

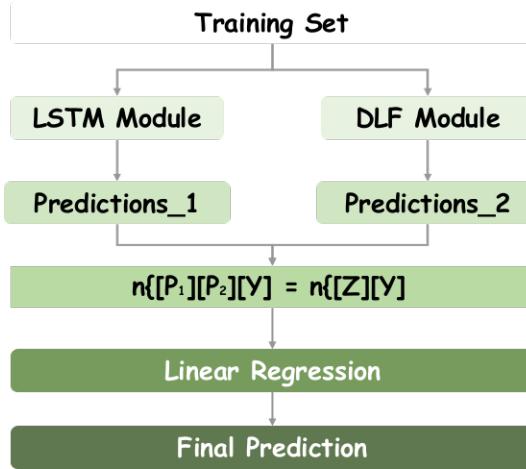


图 2.3 DLF-LSTM 模型架构图

尽管线性回归在处理非线性数据时可能存在一定的局限性，因为它本质上是线性的，但由于我们已经使用了双逻辑函数和 LSTM 对数据进行了非线性处理，线性回归可以有效地作为元学习器来整合这些非线性模型的输出结果。

---

**Algorithm 1** DLF-LSTM
 

---

- 1: Initialize models: Double Logistic Function (DLF) and LSTM
  - 2: Prepare dataset  $D$ , split into  $D_{\text{train}}$  and  $D_{\text{test}}$
  - 3: **Train DLF and LSTM:**
  - 4: **for**  $(x_i, y_i) \in D_{\text{train}}$  **do**
  - 5:     Predict  $y_{\text{DLF}}^i \leftarrow \text{DLF}(x_i)$  and  $y_{\text{LSTM}}^i \leftarrow \text{LSTM}(x_i)$
  - 6:     Store predictions
  - 7: **end for**
  - 8: Combine predictions into new feature set  $X_{\text{meta}}$
  - 9: **Train Linear Regression Meta-Learner:**
  - 10: Train on  $X_{\text{meta}}$  to predict  $y_i$
  - 11: **Evaluate Model on  $D_{\text{test}}$ :**
  - 12: **for**  $(x_i, y_i) \in D_{\text{test}}$  **do**
  - 13:     Predict  $y_{\text{pred}}^i$  and compute accuracy
  - 14: **end for**
  - 15: Output the overall performance
- 

使用“堆叠回归”(Stacked Regression)的方法将 DLF 和 LSTM 融合，能够在一定程度上实现两者的互补，充分发挥各自的优势，从而提高整体模型的预测准确性和鲁棒性。

## 2.5. 模型对比

我们对融合之后的 DLF-LSTM 模型与其他模型及融合之前的模型进行了对比，比较了不同机器学习模型在测试集上的性能，尤其是在误差（MSE、RMSE）和模型解释能力（ $R^2$ ）方面，结果如下图所示。

模型类型	RMSE	$R^2$
线性回归	3,275.755	0.719
LightGBM	25,370.049	-15.846
随机森林	15,950.829	-5.293
DLF-LSTM	824.94	0.981

模型类型	RMSE	$R^2$
DLF-LSTM	824.94	0.981
LSTM	9401.701	-0.321

图 2.4 与原先模型的对比

从图表中可以看出，DLF-LSTM 模型在所有指标上表现最优，具有最低的 MSE 和 RMSE，以及非常高的  $R^2$  值，显示了出色的预测能力和数据拟合度。相比之下，其他模型的表现则不尽如人意。例如，随机森林模型出现了过拟合现象；LightGBM 的  $R^2$  值甚至为负，显示其在测试集上的表现非常差；线性回归模型的准确率不高，难以胜任复杂的预测任务。

对比融合之前的模型，LSTM 的 RMSE 明显高于 DLF-LSTM，表明其在测试集上的预测误差较大。此外，LSTM 的  $R^2$  值为负，表示其在测试集上的表现甚至比基于数据平均值的简单模型还要差，这是典型的过拟合迹象。

总体而言，DLF-LSTM 模型展示了非常低的误差和高度的数据拟合能力（ $R^2$  接近 1），表明其具有优越的性能，能够有效地避免过拟合，并在未见数据上提供可靠的预测。而标准 LSTM 模型则可能需要进一步的参数调整和正则化以改善其泛化能力。

这个对比结果表明，DLF-LSTM 模型在各种指标上均表现出色，是一个更为优越的特定于少量时序数据的选择，能够在复杂的预测任务中提供更高的准确性和可靠性。



# Chapter 3

## 背景

### 3.1. 两会中的乡村发展

在中国这片辽阔的土地上，城市的霓虹灯光下掩映着农村的静谧星空。改革开放以来，中国经历了翻天覆地的变化。城市化与工业化的巨轮迅速推进，城乡人口结构发生巨大改变，同时也带来一系列“三农”问题。在这样的背景下，乡村振兴战略的提出成为解决社会主要矛盾的必然需求，是推动城乡协调发展的关键举措。

党的十四届全国人大二次会议中多次提到，推进乡村全面振兴是新时代新征程“三农”工作的总抓手，要锚定建设农业强国目标，学习运用“千万工程”经验，以加快农业农村现代化更好推进中国式现代化建设。要坚持农业农村优先发展，加快现代农业建设，全方位夯实粮食安全根基，多途径促进农民收入较快增长。要坚持农民主体地位，大力培养乡村人才，吸引各类人才投身乡村振兴。要深入推进农村生态文明建设，加快发展方式绿色转型，建设宜居宜业和美乡村。

本文通过爬虫技术梳理了全国两会报告中涉及乡村发展的所有内容，旨在彰显国家对乡村振兴战略的高度重视。

推动城乡融合和区域协调发展  
智能化，让生活更美好。打好乡村全面振兴漂亮仗。

推进绿色低碳发展 大力优化经济布局  
扎实推进乡村全面振兴 绿色化转型  
今年要把促进农业转移人口市民化放在突出位置来抓  
从两会看中国经济 推动传统产业高端化  
三次 政府工作报告中的绿色低碳发展

图 3.1 两会爬虫文本看乡村发展必要性

自党的十九大报告首次提出实施乡村振兴战略以来，中国在“三农”工作上已取得了一定成就，本文收集并整理了近年来的相关数据，致力于分析农村发展历程。

## 3.2. 角度分析

在了解到国家高度重视乡村振兴战略的基础上，本文将站在乡村发展的视角上深入探讨这一战略。通过整合《全国乡村产业发展规划（2020-2025年）》的相关文本并筛选出关键词汇，以词云的形式展现本文分析乡村发展的核心内容。



图 3.2 全国乡村产业发展规划词云图

在词云中突显出诸多关键词，如“农产品”、“休闲”、“创新创业”、“建设”、“旅游”、“农民”和“就业”等，体现出乡村振兴战略的多维焦点，也为本文后续的论述指明方向。因此，本文将以这些关键词作为基础，解析它们在乡村振兴过程中的具体含义与作用，进而深化对该战略全局性和战术性贯彻落实的认识。



## Chapter 4

# 粮食安全问题

习近平总书记指出，粮食安全是“国之大者”，耕地是粮食生产的命根子，确保 18 亿亩耕地红线决不突破。

在全球化的背景下，粮食安全问题成为国际社会关注的焦点。作为世界上人口数量位居前列的国家，中国在粮食安全方面的表现不仅关乎其经济发展和社会稳定，还对维护全球粮食安全体系发挥着关键作用。而农村地区，作为粮食生产的主要基地，承担着维护国家粮食安全的重要责任。

### 4.1. 农作物播种面积

粮食安全的基础在很大程度上依赖于稳定且充足的播种面积。播种面积不仅直接影响到粮食的总产量，也关系到农业的可持续性和对环境变化的适应能力。本文收集到了 1990-2022 年的农作物总播种面积。为直观分析这一时期内播种面积的波动和趋势，本文计算出播种面积的平均值，并利用玫瑰图展示各年度播种面积相对于平均值的变化，如下图所示：

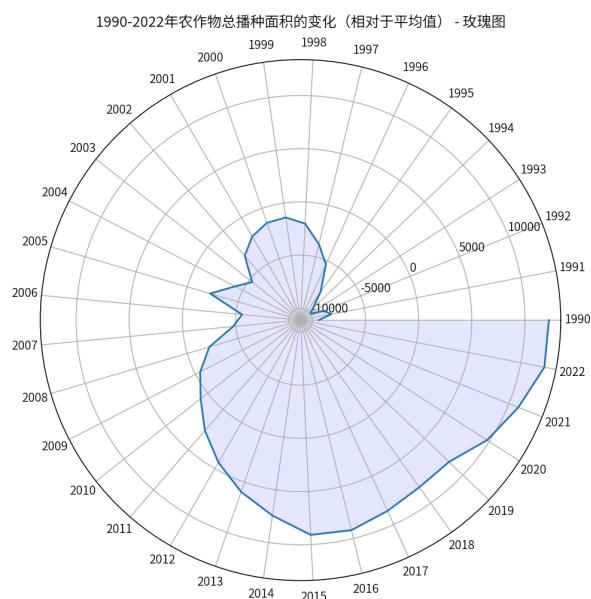


图 4.1 1990-2022 年农作物总播种面积变化图

从图中可以看出，大多数花瓣沿着径向向外扩展，这表明随着时间的推进，播种面积呈现出增长趋势。这种增长通常指示了农业用地的扩张，对粮食安全来说是一个积极的信号，因为这意味着有潜力生产更多的粮食来满足日益增长的人口需求。同时，这样的增长也应当与生态和可持续发展原则相协调。

## 4.2. 主要粮食产量

通过阅读农业农村部门文件农办科〔2023〕15号以及其他相关文献，可以了解到中国的农业生产主要集中在小麦、稻谷、大豆和玉米等主食农作物上。这些作物的生产状况对我国的粮食安全至关重要。

同样，本文收集了1990年到2022年间的粮食产量数据。为了更清晰地展示近年来的发展趋势，本文截取了2005年至2022年间的数据绘制了折线图，如下图所示：

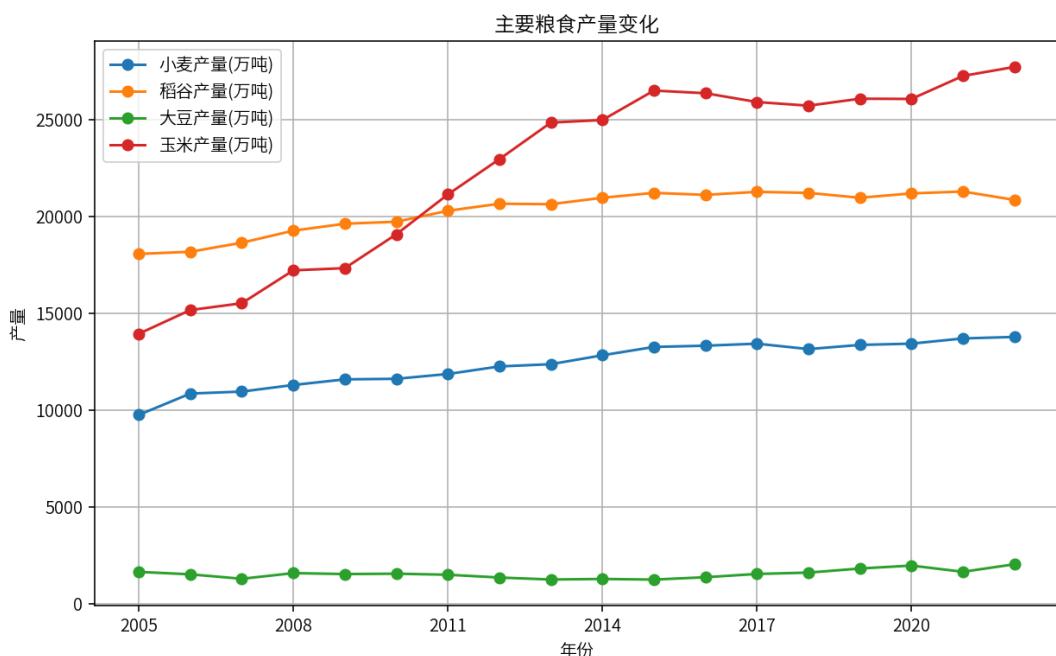


图 4.2 主要粮食产量变化

此折线图清晰地呈现了自2005年以来中国主要粮食作物的产量变化。显著的是，水稻、小麦和玉米三者的产量逐年攀升，这一趋势不仅标志着农业技术进步和生产效率的提高，也反映了中国乡村振兴策略成功增强了农村地区的农业生产能力。鉴于这些作物在中国饮食中占据基础性地位，它们产量的增加直接增强了国家粮食安全保障。

此外，在十四届全国人民代表大会第二次会议上，国务院提出了我国粮食库存消费比的最新数据，显示我国的粮食库存消费比远高于联合国粮农组织建议的17%-18%的国际安全水平。这一数据不仅展示了我国在粮食安全领域所取得的显著成就，也反映了国家对于保障粮食安全的高度重视和扎实努力。



# Chapter 5

## 科技进步

《“十四五”推进农业农村现代化规划》提出 [1]，计划到 2025 年，农业基础更加稳固，乡村振兴战略全面推进，农业农村现代化取得重要进展。科技的进步显著推动了乡村发展，不仅改变了农村的面貌，也为乡村振兴注入了新的活力 [2]。以下是十四五规划的规划目标图：

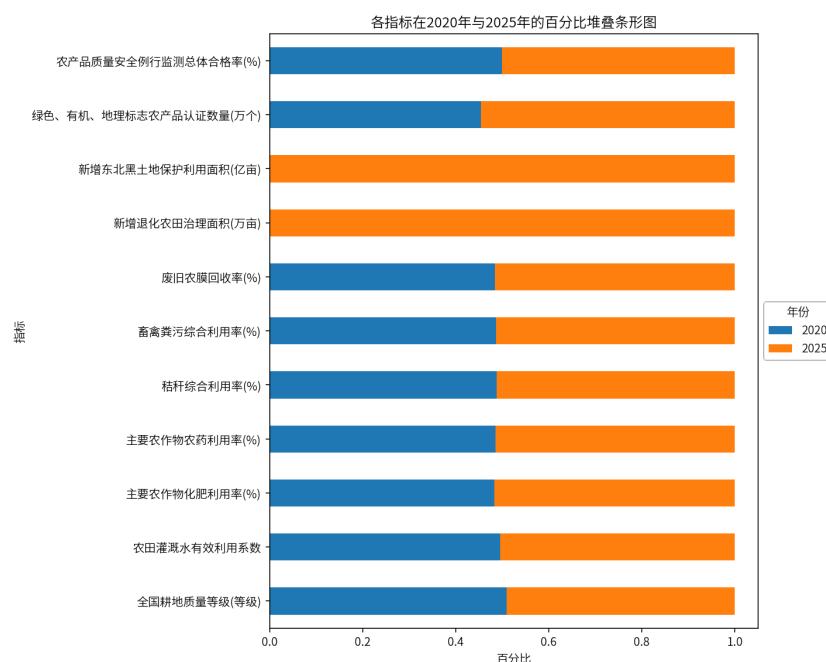


图 5.1 十四五规划流程目标图

### 5.1. 科技对乡村农业的发展的影响

#### 5.1.1. 科技主要贡献

农业现代化是实现可持续发展的关键，而农业科技现代化则是推动农业现代化的核心。本文综合收集并分析了农业科技在全国范围内的核心指标，涵盖农业科技进步贡献率、化肥与农药的有效利用率，以及县域数字农业发展的现状。依据这些关键指标，搜集了相应的数据，并据此绘制了热力图，如下图所示：



图 5.2 全国农业主要科技贡献相关指标的热力图

由图可知，科技在农业中的贡献逐年上升，体现在农业科技进步贡献率的持续增长。这不仅意味着科技创新在提高农业生产力方面起到了关键作用，而且反映出乡村经济正逐渐向技术密集型发展模式转变。

化肥和农药的有效利用率的增加，指向了农业生产过程中资源利用效率的提高。县域数字农业农村发展总体水平的提升也标志着信息技术在乡村振兴中的日益重要性，表明数字化转型在助力农业现代化和乡村经济发展中发挥着增效作用。

### 5.1.2. 数字农业渗透率

随后，本文特别关注了数字农业渗透率这一关键指标，以评估技术如何在实际农业生产中得到应用并推动乡村经济的发展，绘制出如下的条形图 [3]：

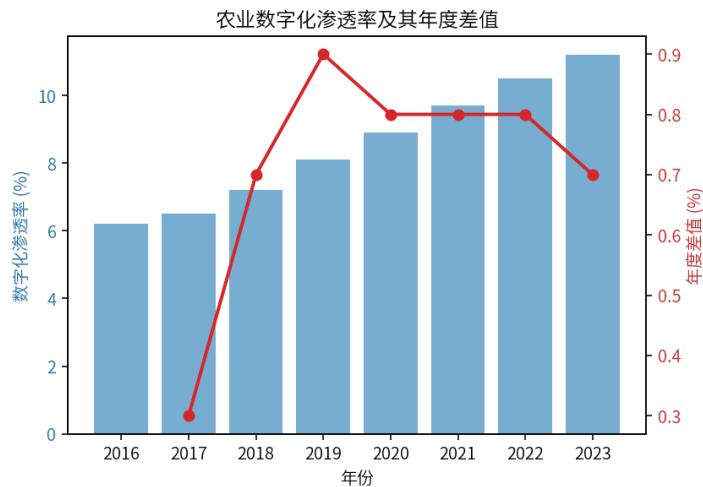


图 5.3 数字农业渗透率折线图

根据图表，可以看到农业数字化渗透率从 2016 年的 0.062 逐年增长，到 2023 年

达到了 0.112。从数字农业渗透率的变化趋势来看，数字化在农业中的应用呈现出逐年增长的态势，说明越来越多的农业生产活动正在通过数字技术进行优化和改进。

在本文中，采用了 DLF-LSTM 模型来对农业数字化渗透率进行建模和预测。在进行了详尽的测试之后，发现当学习率设置为 0.004 且经过 100 次迭代后，模型的损失率降至 0.00138，并逐渐趋于稳定，这表明该模型达到了较高的预测精度。下面展示的图形描绘了模型训练过程中损失率的变化趋势，从中可以直观地观察到模型优化的效果和损失率随迭代次数的下降情况。

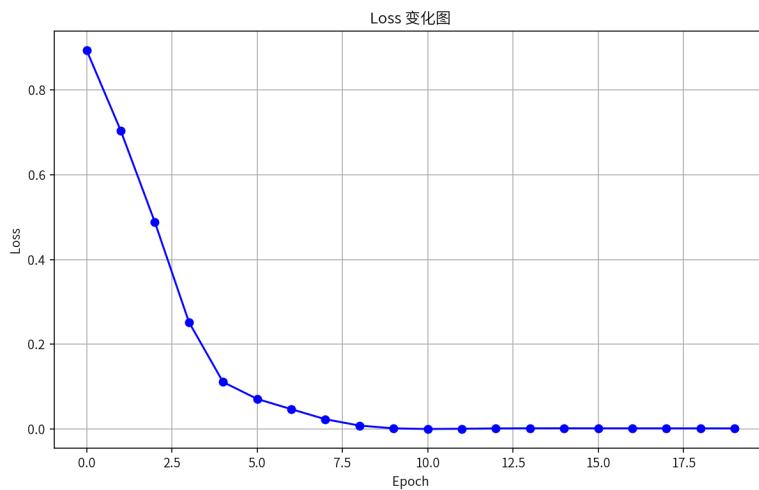


图 5.4 loss 损失图

预测数据如下表：

表 5.1 农业渗透率预测表

年份	预测农业渗透化率（万人）
2024	0.12
2025	0.13
2026	0.14
2027	0.15

从预测数值来看，农业数字化渗透率在未来几年将保持稳步增长的态势。从 2021 年的 0.12 逐步增长到 2027 年的 0.15，虽然整体呈现出向上的趋势，但增长幅度不是很大，且普及程度尚未达到预期的高水平。因此，可能需要进一步加大力度，推动数字技术在农业和乡村地区的广泛应用，以实现数字农业的全面渗透，带动乡村经济的全面发展。

### 5.1.3. 农业机械总动力

在当前农业现代化的大背景下，农业机械化的推进尤显重要。本文借助国家统计局提供的权威数据，深入分析了我国农业机械化的发展态势。农业机械总动力，作为衡量农业机械化水平的重要指标，反映了全部农业机械动力的综合实力。绘制如下条形图进行分析：

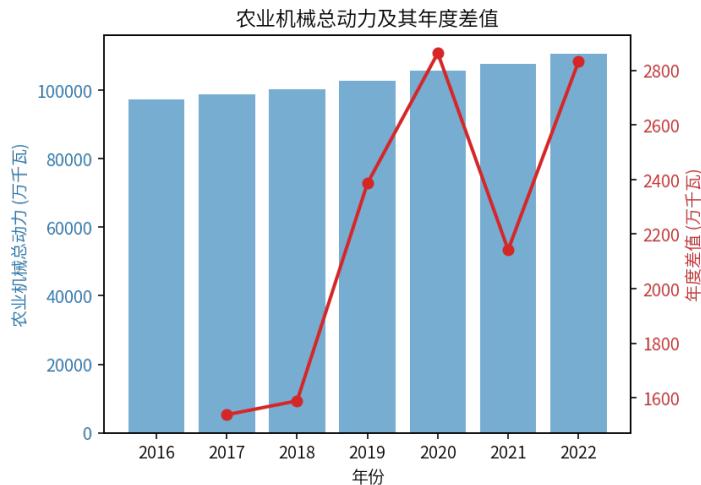


图 5.5 农业机械总动力条形图

数据显示，我国农业机械总动力呈持续上升趋势。截至 2022 年，该指标达到了 110597.19 万千瓦，较去年增长了 2800 多万千瓦，展现了我国农业机械化快速发展的势头。

利用时间序列预测模型，进一步预测了未来的发展趋势。模型显示，若当前发展势头保持不变，农业机械总动力有望在未来几年内达到 114378.54 万千瓦。农业机械总动力的逐年提升，不仅彰显了国家在农机装备产业转型升级方面的坚定决心，也体现了农业农村部在推进“十四五”全国农业机械化发展规划中的积极作用 [4]。

表 5.2 农业机械化预测表

年份	未来农业机械总动力 (万千瓦)
2023	114378.54
2024	117931.49
2025	121980.57

农业机械化的持续发展是国家现代化农业战略的重要组成部分。随着政府对农业机械化的持续投入和政策支持，有理由相信，我国农业机械化的未来将更加光明，为实现乡村振兴战略贡献更大的力量。

## 5.2. 科技对乡村居民生活的影响

随着我国乡村振兴战略的深入实施，科技创新力量不断渗透到乡村生活的各个方面，显著改变了广大农村居民的日常生活习惯和经济社会发展面貌。近年来，国家积极推动互联网基础设施向农村延伸，加快缩小城乡数字鸿沟，这一举措成效显著。

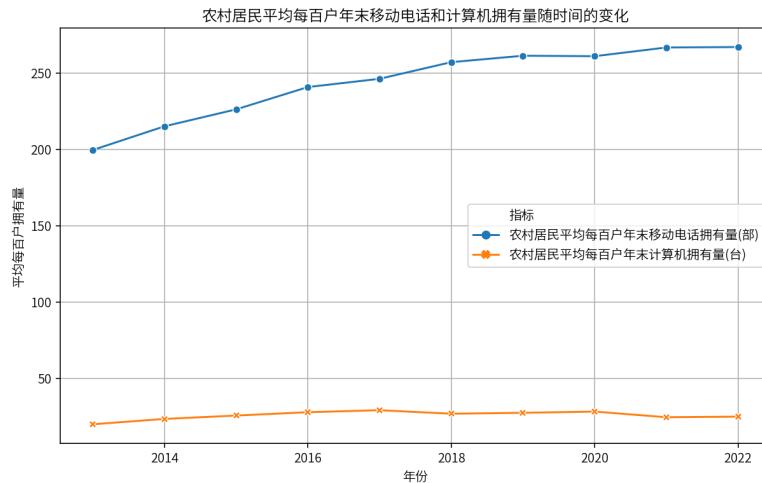


图 5.6 电子设备使用量序列图

本文通过绘制时间序列的折线图，进而揭示了农村居民对电脑和手机这两类主要电子设备的使用趋势。随着互联网技术的广泛普及，可以观察到农村居民的手机拥有量呈现出显著的增长趋势，反映了移动设备在日常生活中的重要地位。相对而言，电脑的使用量在这段时间内变化不大，暗示着农村地区对于电脑的需求和应用场景仍然有限，或许是由于农村地区对高科技设备的适应性和接受度有待提高。

通过对城乡居民互联网使用状况的数据分析可以发现，无论是城市还是乡村，互联网普及率均呈现出稳步上升的趋势。特别是在过去几年间，随着移动互联网技术的普及和网络资费的下调，越来越多的农村居民得以接入互联网，享受到便捷高效的数字化服务 [5]。

表 5.3 2020-2022 城乡居民互联网使用量预测

年份	农村居民互联网使用量预测	城镇居民互联网使用量预测
2020	63.680962	77.475733
2022	82.730215	98.743864

本数据缺失 2020 到 2022 的数据，因此使用指数平滑法进行预测。结合趋势推测，当前城乡互联网用户的比例可能接近于 5:4 这样一个水平，预示着乡村互联网市场潜力巨大，并有望在未来进一步激活农村经济，助力乡村振兴战略目标的实现。

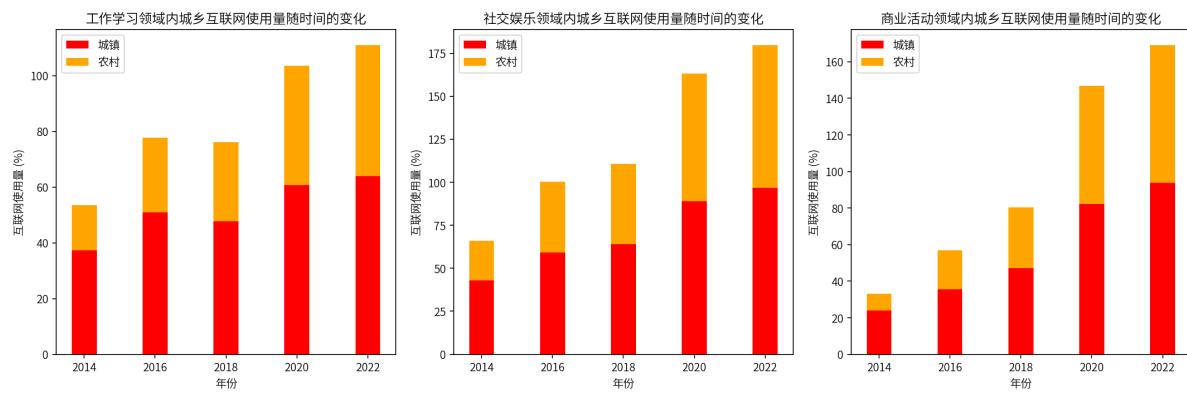


图 5.7 城乡互联网使用量随时间变化对比图

数据显示，城乡间的互联网使用量差距正在逐年缩减。农村地区的互联网普及率不断提升，尤其是在商务交易、网络支付、信息获取等方面的应用率增长明显，这表明农村居民正加速融入数字经济的大潮之中。例如，农村网民通过电商平台进行农产品销售、利用网络平台开展技能培训、借助移动支付实现无现金交易等活动日益普遍。

这些数据共同映射了中国乡村经济在科技推动下的积极变化，揭示了科技进步是乡村振兴策略的关键驱动力。



# Chapter 6

## 乡村建设

《乡村建设行动实施方案》指出：乡村建设是实施乡村振兴战略的重要任务，也是国家现代化建设的重要内容。

乡村基础设施的建设不仅是提升乡村生活品质的关键，更是吸引人才和促进人口回流的重要举措。此外乡村地区的文化和自然环境也是其独特的吸引力所在。

### 6.1. 基础设施建设

在具体的基础设施建设中，为了量化评估进展和成效，本文选择了两个关键性的指标：每万人拥有的卫生床数量和人均养老服务机构数量。这两个指标不仅反映了乡村医疗卫生和养老服务设施的发展水平，也是衡量基础设施建设对提升居民生活质量影响的重要参数。

在收集近十年各地区的卫生床和人均养老服务机构数量数据的过程中 [6]，本文采用年复合增长率（CAGR）这一指标，将每个地区过去十年的数据统一转化，以便进行比较。同时进一步制作了一幅详尽的中国热力图，直观展现了各地区在所考察指标上的表现。

中国地图-乡镇每万人卫生床位数

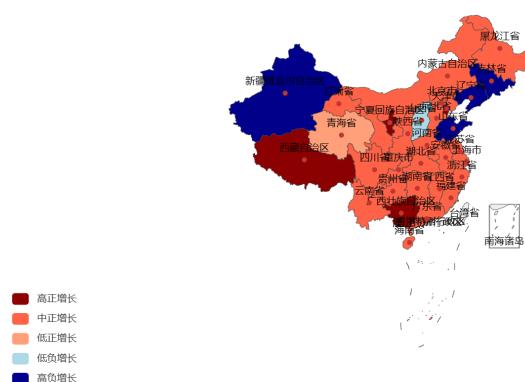


图 6.1 卫生床位数

中国地图-人均养老服务机构数量

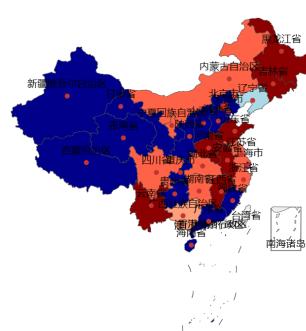


图 6.2 养老服务机构数量

在这两幅图中，红色区域表示正增长区域，蓝色区域表示负增长区域。颜色的深浅代表了增长的强度：红色越深表明正增长程度越高，蓝色越深则意味着负增长程度越显著。

结合胡焕庸线的地理分布来观察，在沿线东部地区，尤其是人口密集的乡镇，无论是卫生床位数还是养老机构数量，普遍呈现出正增长趋势。

然而，胡焕庸线以西的西部地区，尤其是人口较少的农村区域，却出现了卫生床位和养老机构数量的负增长趋势。这种情况在一些经济较发达的地区也有所体现，可能与这些地区的农村人口减少，需求降低有关。

这种差异性的现象揭示了我国农村在基础设施建设方面正处于一个积极变化的发展阶段。东部的积极增长反映了国家对农村地区尤其是人口密集区域的关注和投资的增加，显示出公共服务水平正在逐步提升，以适应居民不断增长的需求。而西部地区则突显了在人口减少的背景下，如何有效地规划和调整基础设施，以持续满足农村居民变化的需求。

## 6.2. 基础环境建设

在推进乡村基础设施建设的过程中，必须同步加强对乡村自然环境的保护和改善，这对于实现乡村振兴和可持续发展具有至关重要的作用。

本文对乡村环境进行了系统的量化分析，收集并研究了 2015 至 2022 年期间的关键环境指标数据 [7]。这些数据包括村庄的供水情况、排水管道系统、人均绿地面积、环卫车辆的数量和使用情况以及公共厕所的配置和状况。通过计算连续两年之间的数据差值来确定每年的变化量，绘制出变化量与年份的关系图如下图所示 [8]：

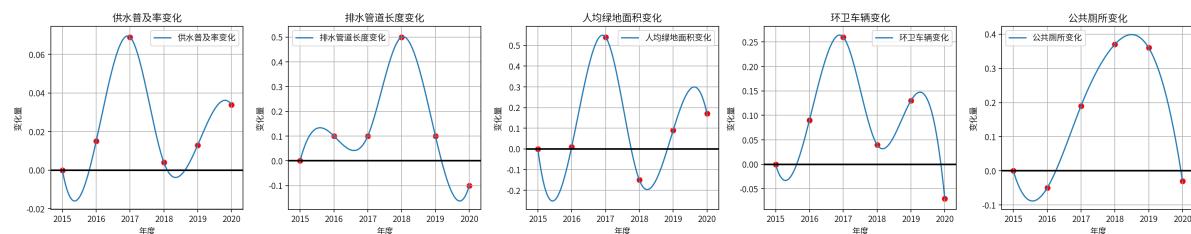


图 6.3 基础环境改善图

图表展示了多个指标随时间变化的趋势，其中大部分曲线持续处于零点线以上。这表明，尽管各指标呈现出明显的波动，总体上还是保持了增长态势。这一现象彰显了国家对基础设施与环境改善的坚定承诺。

## 6.3. 生态旅游建设

乡村环境的改善已成为提升居民生活质量和健康水平的关键因素，同时也开辟了乡村经济转型升级的新路径。

环境质量的提升不仅吸引了生态旅游和绿色农业等绿色经济产业，也促成了新的经济增长点和增加了就业机会，为乡村带来了新的活力。本文深入分析了过去近十年

来乡村旅游接待游客数量及其所带来的旅游总收入 [9]，绘制出如下图所示的柱状图：

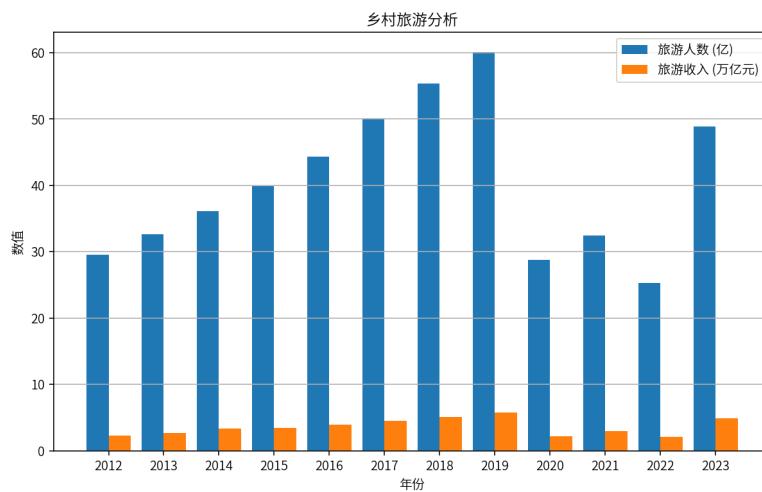


图 6.4 近十年乡村旅游人数与旅游收入

初期数据显示，乡村旅游人数和收入都呈现出稳步增长的态势，反映了人们对逃离城市喧嚣、探索自然和体验乡村生活愈发浓厚的兴趣。然而，2020 年新冠疫情的爆发对我国旅游业造成了前所未有的打击，乡村旅游也未能幸免，其人次和收入均出现了急剧下滑。

随着疫情的结束，乡村旅游业开始呈现出复苏的势头，人们逐渐回归到了疫情前的生活和休闲模式。乡村旅游经历了一段迅猛的增长期，这种增长不仅归功于疫情的消退，乡村基础设施和环境设施的改善也在促进乡村旅游发展中发挥了至关重要的作用。



# Chapter 7

## 农民生活水平

农民作为农村的主体，不仅是农业生产的直接参与者，也是乡村社会结构的基础。因此，农民的生活水平是衡量农村经济和社会发展的重要指标，农民生活条件的改善直接影响到农村的可持续发展。

### 7.1. 农村就业

#### 7.1.1. 就业结构改善

随着经济全球化的深入发展和内需市场的不断扩大，科学技术取得显著进步，农民对于提高生活质量的追求日益增强。同时，农民对新知识和新技术的接受度也显著提高，城乡之间的互动更加频繁和深入。这些因素共同作用，推动了农村产业结构的多元化发展。

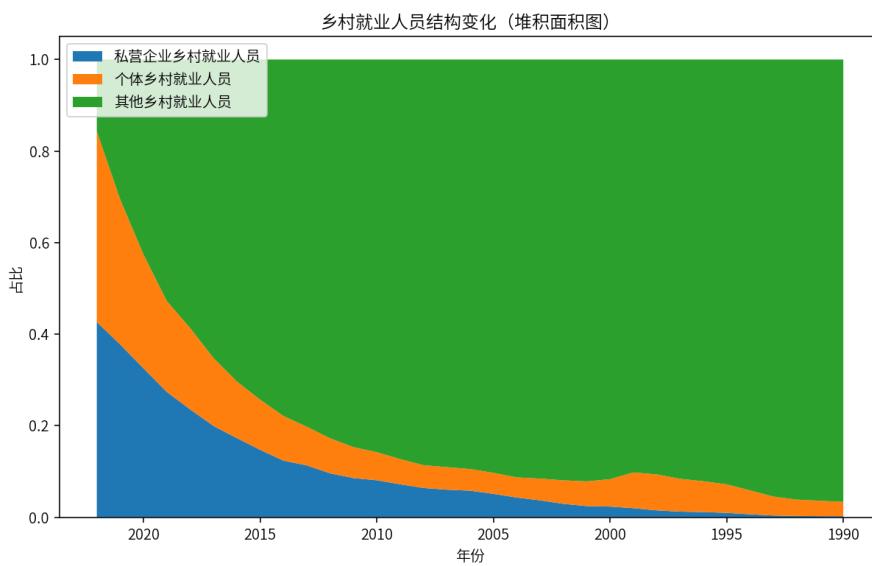


图 7.1 农村产业结构变化

从图中可以观察到，私营企业和个体户在就业人员中的比例呈现出上升趋势，这表明农村就业结构正在经历显著的变革。相对于过去农业的主导地位，农村经济的多

元化趋势愈发明显。如农业深加工、乡村旅游、农家乐等新兴产业，它们在增加就业机会的同时，也极大地丰富了农村的产业内容。

### 7.1.2. 就业人员预测

本文综合分析了 1990 年至 2022 年间的城乡就业人员数据，以探索和呈现过去几十年中城乡就业结构的变迁。在数据收集的过程中，发现某些年份的数据存在缺失（标记为 NaN 值），可能会影响对就业趋势的全面理解和预测。为了弥补这一空缺并提高数据完整性，于是考虑采用 DLF-LSTM 模型来进行预测。

首先，模型在评估状态下预测了空缺三年的私营企业的就业人数，并将这些预测数值转换回原始规模以制作预测图表。这张图表显示了历史数据与模型预测的对比。然后，使用同一模型对整个历史数据集进行了拟合，同样地将预测结果转换回原始规模。通过这种方式制作了第二张图表，展示了模型对已知历史数据的拟合程度。

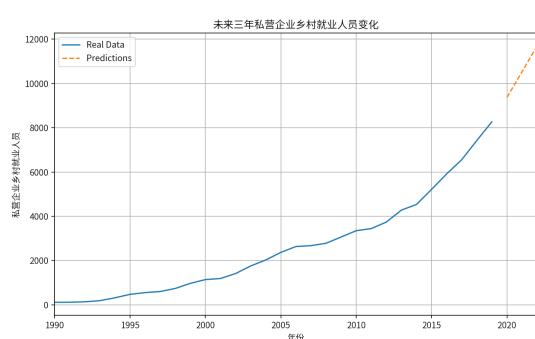


图 7.2 未来三年私营企业乡村就业人员变化图

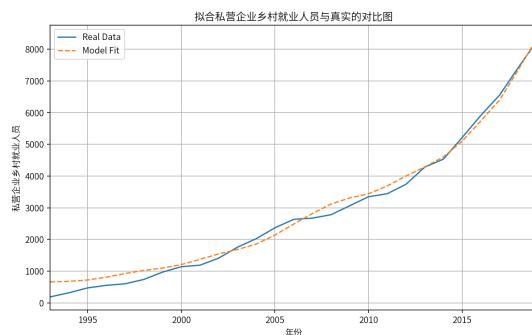


图 7.3 拟合私营企业乡村就业人员与真实的对比图

根据所呈现的图表分析，可以观察到该模型具有较高的拟合度，并且能够有效地预测趋势。这表明该模型对于描绘乡村私营企业就业人数的变化具有可靠性。利用此模型，成功估算了三年内私营企业乡村就业人员的空缺数据并预测了后三年的结果，详细结果如下表所示。

表 7.1 私营企业乡村就业人员数据

(a) 补全空缺三年私营企业乡村就业人员		(b) 预测三年私营企业乡村就业人员	
年份	私营企业乡村就业人员	年份	私营企业乡村就业人员
2020	9499.59	2023	13408.71
2021	10732.58	2024	14651.86
2022	12013.71	2025	15764.10

同样地，本文也对个体户乡村就业人员的数据集进行了处理，并生成了相应的图表和数据。

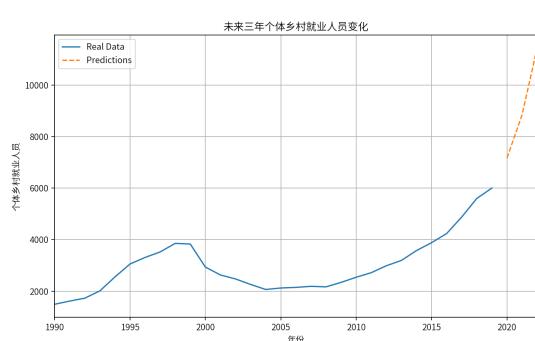


图 7.4 未来三年个体户乡村就业人员变化图

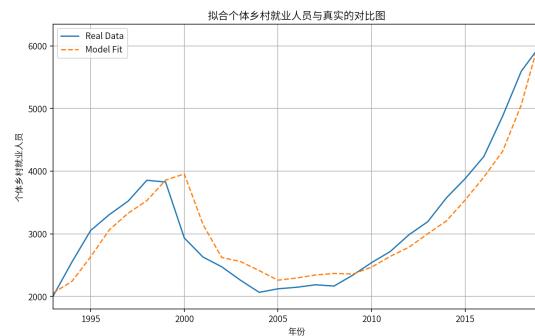


图 7.5 拟合个体户乡村就业人员与真实的对比图

表 7.2 个体户乡村就业人员数据

(a) 补全空缺个体户乡村就业人员

年份	个体户乡村就业人员
2020	7228.06
2021	8956.09
2022	11552.06

(b) 预测后三年个体户乡村就业人员

年份	个体户乡村就业人员
2023	14889.481
2024	17813.33
2025	19633.70

据数据预测，个体户乡村就业人员和私营企业乡村就业人员将继续实现较大增长。这表明，农业就业的核心理论正在从传统范式转变为以个体户和私营企业为主导。

采用与前述相同的方法，本文利用 DLF-LSTM 模型对城乡就业人员数据进行了预测。经过模型的训练与验证，得到了预测未来十年城乡就业人数变化的图表和数据。

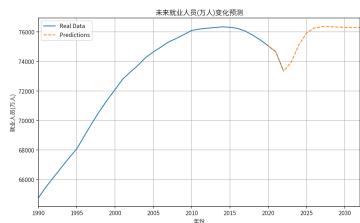


图 7.6 未来就业人员变化预测

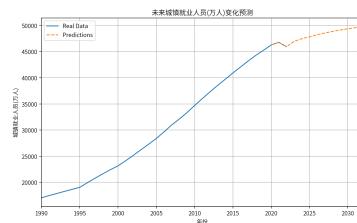


图 7.7 城镇就业人员变化预测

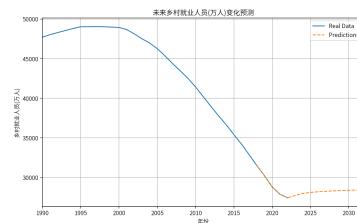


图 7.8 乡村就业人员变化预测

图表显示总就业人数在经历了一个上升趋势之后，趋于平稳并略有波动。预测显示未来会有一定的波动，但总体维持在一个相对平稳的水平。这可能意味着就业市场在达到一定规模后，将进入一个较为成熟的阶段，增长速度放缓。

预计城镇就业人员数量将继续呈现上升趋势，不过增长速度逐渐放缓，显示出一个逐渐稳定的趋势。这表明城镇就业市场的发展已经逐渐趋向成熟，未来的增长可能更多地受到经济周期和政策调控的影响。

表 7.3 预测就业人员变化

表 7.4 城镇就业人员变化预测

表 7.5 乡村就业人员变化预测

年份	预测就业人员 (万人)	年份	预测城镇就业人员 (万俟份)	预测乡村就业人员 (万人)	
2020	73939.57	2020	46958.44	2018	27709.54
2021	75116.24	2021	47466.87	2019	27967.95
2022	75921.16	2022	47811.33	2020	28101.58
2023	76266.91	2023	48213.55	2021	28183.98

农村就业人员数的预测在初始阶段显示下降趋势，随后转为平稳甚至上升趋势。这可能反映了随着农业自动化和工业化的推进，农村劳动力需求减少，但随着政策支持和农村经济结构的逐步调整，农村就业人数将趋于稳定，或呈现增长趋势。

## 7.2. 可支配收入

农民生活水平的提升建立在农民可支配收入的基础上，可支配收入直接影响到农民的消费能力。只有当农民拥有足够的可支配收入时，他们才能够改善自身及家庭的生活质量，满足日益增长的物质和文化需求，进而推动农村社会的整体进步与繁荣。

本文详细汇编并分析了从 1999 年至 2022 年间，中国农村地区居民的可支配收入与消费支出的全面数据集。研究的核心在于运用斯皮尔曼等级相关系数这一强有力的非参数统计工具，该方法无需假设数据服从特定分布，即可有效评估两个变量间是否存在单调递增或递减的关系。通过计算如下公式

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad ^1 \quad (7.1)$$

我们深入探讨了收入与支出变量间潜在的关联模式。分析结果显示，相关性图表清晰地揭示了农村居民可支配收入与其消费支出之间存在极为紧密且显著的相关性趋势。

<sup>1</sup>其中  $\rho$  表征相关程度， $d_i$  为每一对观测值在各自排序后之差，而  $n$  代表样本总量



(a) 收入数据的相关性分析

(b) 支出数据的相关性分析

图 7.9 斯皮尔曼等级相关系数图表

为进一步深化理解，本研究采取了主成分分析（PCA）技术，这是一种强大的降维手段，旨在从复杂的数据集中提炼出最具解释力的综合指标。PCA 通过将原始变量转换到新的坐标系统中，即主成分方向上，以最能解释数据方差的方式重新表达数据，其结果不仅依赖于数据自身的属性，还与预处理步骤如数据的中心化和标准化密切相关。所获得的 PCA 图表生动展现了收入与支出数据在新维度上的结构特征，有助于揭示隐藏的模式和减少冗余信息。

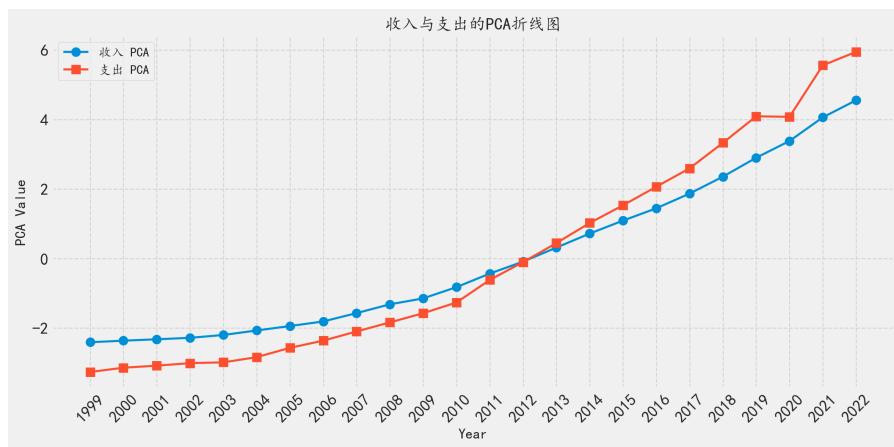


图 7.10 主成分分析图表

在此基础上，我们采用牛顿多元法执行非线性回归分析，以模型  $a * x^2 + b * x + c$  描述农村居民收入与消费支出随时间变化的动态规律，成功拟合了数据曲线。这一发现不仅加深了我们对农村经济活动周期性和长期趋势的理解，也为政策制定者提供了宝贵的量化依据。

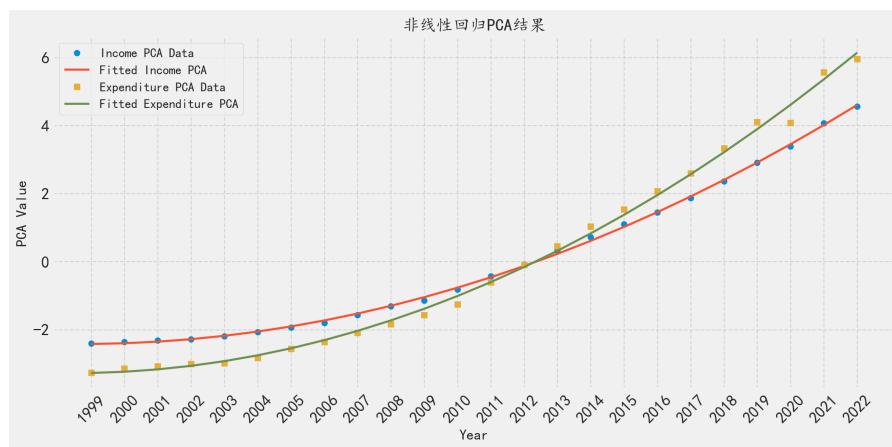


图 7.11 非线性回归 PCA 结果

最后，通过对平均收入与平均消费支出的年度折线图，研究揭示了一个令人关注的趋势：两者间的差距随着时间推移而逐渐扩大，这凸显了农村经济发展中人民生活愈加富裕，表现了制定精准有效的政策措施，促进农村经济的均衡发展和提升农村居民的生活水平获得卓越成效。

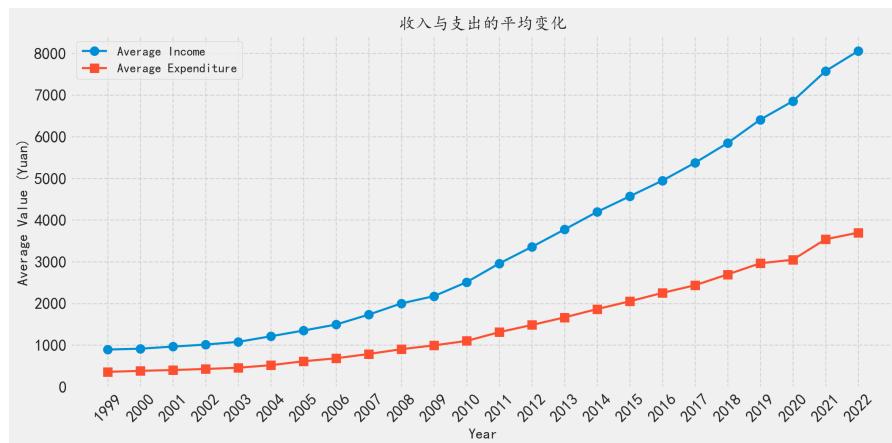


图 7.12 收入支出平均年化折线图

## 7.3. 基本生活

在诸多因素如可支配收入的增加、科技的迅速进步以及政策的积极推动下，乡村居民的生活质量无疑正处于不断提升的轨迹之上。为了深入分析这一变化，我们采集并利用了关于主要交通工具和基础家用电器的拥有量数据，对比了过去十年与当前情况下的数据绘制了雷达图：

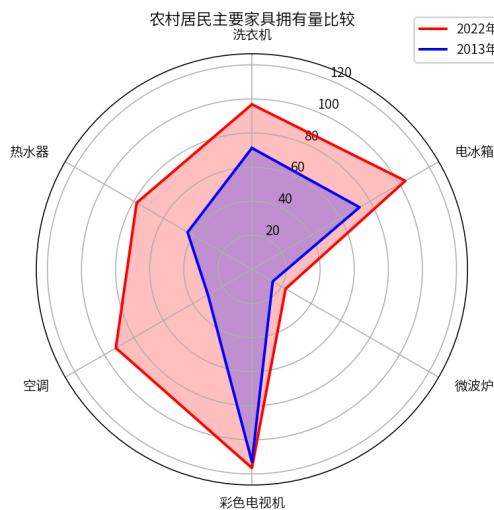


图 7.13 农村居民主要家具拥有量比较

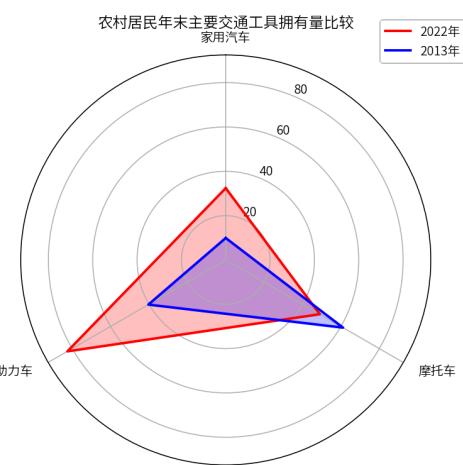


图 7.14 农村居民年末主要交通工具拥有量比较

通过细致构建的雷达图，得以直接洞察到洗衣机、电冰箱、微波炉、彩色电视机、空调以及热水器等基础家电的拥有量显著增加，这一变化凸显了乡村居民生活水平的提升和生活质量的改善。交通工具中，除摩托车受到政策变动的影响出现了下降趋势，家用汽车和电动助力车实现了显著增长，这不仅反映了乡村地区交通工具使用模式的变化，也突显了现代化进程中对于便捷交通方式的日益增长需求。

## 7.4. 幸福指数

根据马斯洛需求层次理论，人的需求从低到高分为生理需求、安全需求、社会需求、尊重需求和自我实现需求。生活水平的提升不应局限于物质层面的丰富，更应包括精神层面的满足。

其中，恩格尔系数，即家庭食品支出占总消费支出的比重，是反映居民生活水平的重要经济指标。该系数越低，表明家庭在食品上的支出占比越小，一般而言，家庭的生活水平越高。

本文搜集了自改革开放以来至今的全国居民（包括城镇和乡村居民）的恩格尔系数数据 [10]，并据此绘制了反映全国总体以及农村居民恩格尔系数变化的折线图，如下图所示。

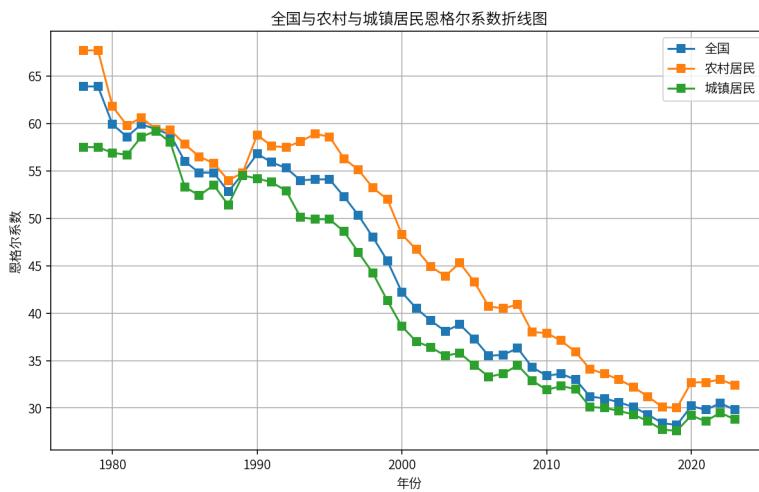


图 7.15 农村人民恩格尔系数折线图

图7.15清楚地揭示了过去数十年间我国居民消费结构的显著变化。全国居民的恩格尔系数呈现下降趋势，反映了随着经济的持续发展和人民生活水平的逐步提高，居民的消费结构正在发生深刻变化。这不仅体现了居民生活质量的整体提升，也反映了消费模式的多元化和层次的提升。

特别值得关注的是，农村居民恩格尔系数的显著下降，凸显了农村居民消费结构的重大转变。农村居民的消费不再局限于基本的食品支出，而是更多地向教育、医疗、娱乐等领域扩展。此外，城乡居民恩格尔系数之间的差异也在逐渐减小，这一现象不仅显示了城乡居民生活水平的逐步接近，也反映了我国在推动城乡一体化发展、缩小城乡发展差距方面取得的积极进展。



## Chapter 8

### 城乡人口结构

下图展示了全球人口结构随年复合增长率变化的情况。从图中可以得出两个主要观察结果：首先，全球发达国家的年复合增长率接近于零，而发展中国家的年复合增长率通常在 1 至 3 之间，非洲等不发达国家的年复合增长率则常超过 5% 且波动较大。其次，像中国这样的发展中国家在图中显示为深绿色，与发达国家形成对比，显示出其巨大的发展潜力和尚未稳定的人口结构比例。因此，发展农村地区成为当前的主要发展目标之一。

全球乡村人口热力图

CAGR

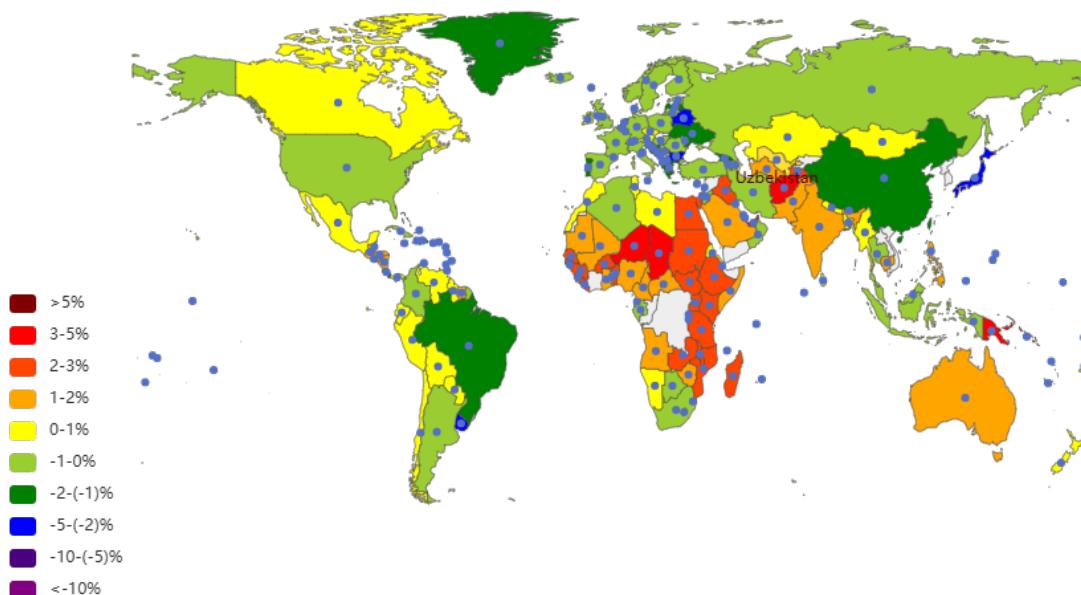


图 8.1 全球乡村人口热力图

自改革开放以来，中国农村面临着劳动力向城市大量流动的现象，这一趋势引发了乡村人口老龄化、劳动力流失等问题。为应对这些挑战，国家提出乡村振兴战略，旨在通过提高农业效率、改善基础设施和公共服务，以此来提升农村生活水平，从而促进农村经济社会全面发展。

随着乡村振兴战略的实施，乡村人口结构出现新变化。我们将利用 1990 年至 2022 年间每隔七年的数据，绘制一系列饼图来展示中国城镇人口和乡村人口的比例。

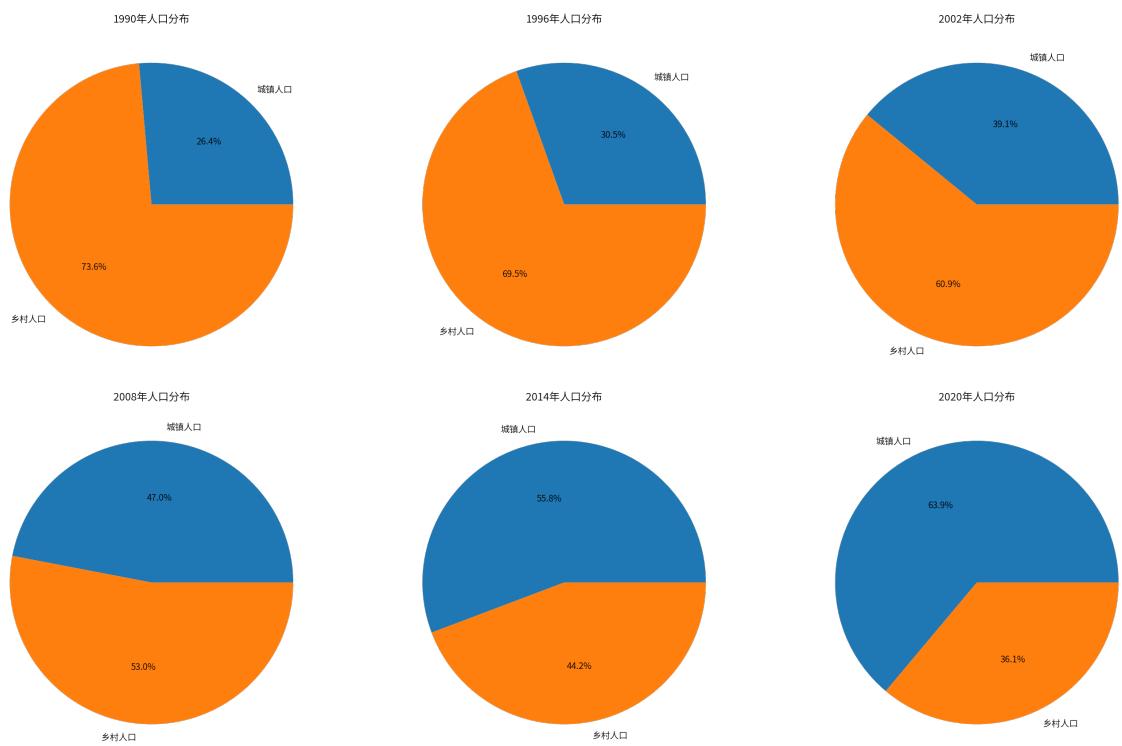


图 8.2 城乡人口占比图

图表显示，城镇人口占比逐年上升，而乡村人口占比相应地逐年下降。这一趋势与改革开放以来中国城市化进程的加快密切相关。

农村地区的人口流失进一步凸显了农村发展的迫切需求。为了缓解这一趋势，必须加强农村建设和政策吸引，提供更多的就业机会，改善教育和医疗条件，以及提高农民的生活水平，为农村居民创造更有吸引力的生活和工作环境，减缓乡村人口的流失速度，促进城乡协调发展。正是在这样的历史背景下，乡村振兴战略应运而生，成为了中国特色社会主义新时代的重要篇章。

## 8.1. 城乡人口预测

进一步分析发现，尽管乡村人口比例持续下降，但下降的速度似乎在放缓，显示出比例减少的幅度逐渐减小。这可能反映了近年来中国在推进城乡发展一体化、实施乡村振兴战略等方面所做的努力开始显现成效，乡村人口流失速度有所放缓。

为了验证城乡人口变化趋势的猜想并预测未来发展，本文采用了 DLF-LSTM 模型预测接下来五年内城乡人口的变化趋势。所得到的预测结果以及模型对历史数据的拟合情况通过下图展示。

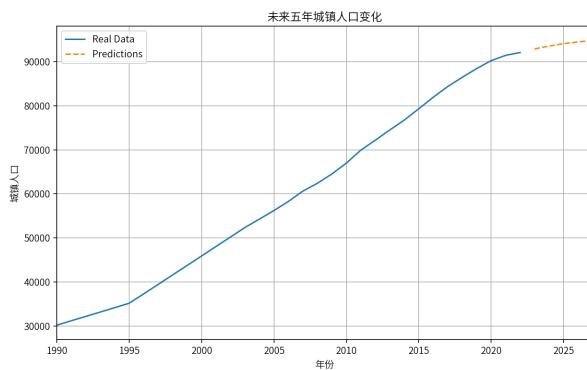


图 8.3 未来五年城镇人口变化

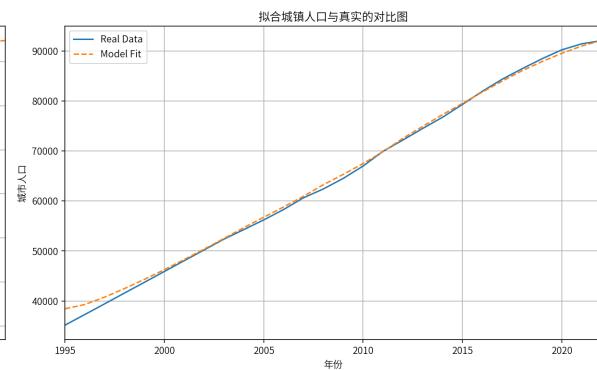
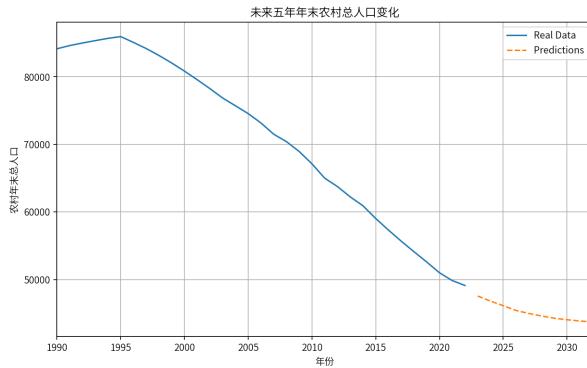


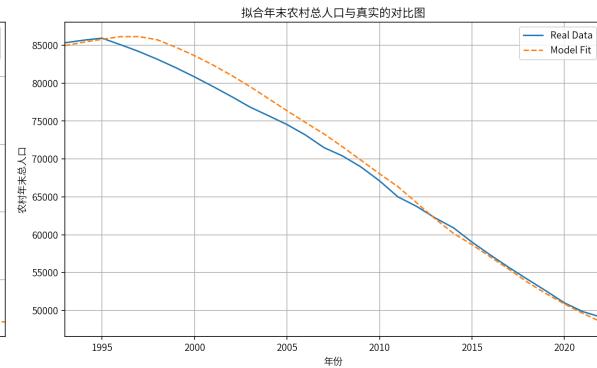
图 8.4 拟合城镇人口与真实的对比图

通过时间序列得知，城镇后几年的趋势可能出现饱和，只有轻微的增长。

接着同样方法分析农村人口的预测，所得结果如下：



(a) 未来五年农村人口变化



(b) 拟合农村人口和真实的对比图

图 8.5 农村人口变化图

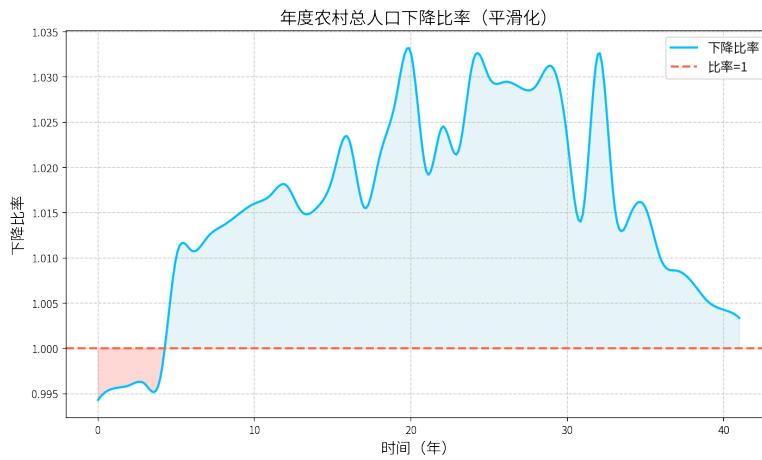
表 8.1 未来城镇人口预测

年份	预测城镇人口 (万人)
2023	92876.58
2024	93542.80
2025	94068.30
2026	94475.73

表 8.2 未来农村总人口预测

年份	预测农村总人口 (万人)
2023	47548.48
2024	46795.32
2025	46118.73
2026	45412.10

随后，本文对每年农村人口变化与前一年的比例进行了汇总，并比较了其与 1 的大小。当这个比例接近 1 时，说明农村人口的减少趋势相对较小，而当这个比例远离 1 时，则意味着变化幅度较大。



三者的结果均显示，农村人口未来的变化趋势或许会受到乡村振兴政策的积极影响，呈现出逐渐趋于平缓甚至可能有所增加的态势。这暗示着在未来一段时间内，农村地区可能会迎来一定程度的人口回流，这可能是乡村振兴战略所带来的积极结果。

这一战略不仅是对中国农村问题的深刻反思，更是对实现城乡和谐发展愿景的坚定追求。在此战略下，农业不再是单一的粮食生产，而是转型为现代化、多元化的产业体系。农村不再是落后的代名词，而是成为充满活力和希望的土地。农民不再是时代的边缘人，而是成为享受现代文明成果的主体。



## Chapter 9

### 结论和展望

在当前的发展阶段，乡村振兴不仅关乎农业、农村和农民的未来，更是国家战略全局的重要组成部分。通过本次数据分析报告，我们得出了一写结论和展望。

必须继续大力稳固粮食安全，这是国家的根本。正如习主席所强调的，“要把中国人的饭碗牢牢端在自己手中”，我们将继续保持粮食生产的稳定性，确保每一粒粮食都能得到妥善利用。

数字农业的发展是提升农业现代化水平的关键。尽管当前数字农业渗透率尚显不足，但我们将加强数字技术在农业中的应用，推动农业生产方式的转变，如《2020 年中国农村数字发展报告》所述，通过数字化转型，实现农业的智能化、精准化。

对于偏远地区，推进基础设施建设是实现乡村振兴的前提。我们将加大投资力度，改善交通、水利、电力等基础设施，为乡村发展提供坚实的物质基础。

乡村旅游业的发展是乡村振兴的一大亮点。我们将依托各地的自然风光和文化遗产，打造特色旅游产品，吸引更多游客，带动乡村经济的多元化发展。

优化产业结构，促进私营企业和个体户的发展，是提高农村经济效益的重要途径。我们将制定更多优惠政策，支持乡村产业的创新和升级，增强乡村经济的内生动力。

为促进人口回流，我们将继续完善公共服务，提高农村居民的生活质量，使乡村成为人们向往的美好家园。如习主席所言，“要让农民在乡村振兴中有更多获得感”，我们将努力实现城乡居民的共同富裕。

乡村振兴战略的实施，需要我们在保障粮食安全、推进数字农业、加强基础设施建设、发展乡村旅游业、优化产业结构和促进人口回流等多个方面下功夫。我们将以实事求是的态度，贯彻落实两会精神和三农政策，为实现乡村全面振兴而不懈努力！



## 参考文献

- [1] “十四五全国农业绿色发展规划”. In: 农业绿色发展 (2019).
- [2] “2023 年农业重大引领性技术及其他相关技术列表”. In: 农办科〔2023〕15 号附件 (2023).
- [3] 中国信息通信研究院政策与经济研究所. “2022 年中国城乡数字包容发展研究报告：数字化助力乡村振兴与共同富裕”. In: 数字农业 (Dec. 2021).
- [4] “中国数字乡村发展报告（2022 年）”. In: 数字农业 (2023).
- [5] 中共中央办公厅国务院办公厅. 中共中央办公厅国务院办公厅印发《数字乡村发展战略纲要》. 2019. URL: [https://www.gov.cn/zhengce/2019-05/16/content\\_5392269.htm](https://www.gov.cn/zhengce/2019-05/16/content_5392269.htm) (visited on 03/25/2024).
- [6] 乡村振兴-生态宜居指标(2012-2020 年). Mar. 2024. URL: <https://www.heywhale.com/mw/dataset/6445342bff981dedde2302d9/file>.
- [7] 严家源. “美好乡土，生态宜居——乡村振兴专题系列报告（环保篇）”. In: 证券研究报告 (Jan. 2022).
- [8] 农业农村部, 国家标准化管理委员会, and 住房和城乡建设部. “乡村振兴标准化行动方案”. In: 农质发〔2023〕5 号 (2023).
- [9] 北京第二外国语学院旅游科学学院. 中国第一本乡村旅游绿皮书《中国乡村旅游发展报告（2022）》在京发布. 2022. URL: [https://www.bisu.edu.cn/art/2022/12/23/art\\_18952\\_307731.html](https://www.bisu.edu.cn/art/2022/12/23/art_18952_307731.html) (visited on 03/25/2024).
- [10] 数据基地. 中国历年恩格尔系数 (1978 年-2021 年). 2022. URL: <https://www.shujujidi.com/caijing/211.html> (visited on 03/25/2024).