

基于主成分分析法和多元回归对山东省农产品冷链物流的实证分析

冯大卫

贵州大学，管理学院，贵州 贵阳

收稿日期：2023年7月1日；录用日期：2023年9月20日；发布日期：2023年9月27日

摘要

随着经济的不断发展，人们对安全、健康的农产品需求越来越多，山东省作为农业大省，拥有丰富的农产品资源，肉类、果类、蔬菜类等农产品的产量位居全国前列。本文以山东省农产品的冷链物流为研究对象，从区域经济发展状况、交通运输水平、国内外市场供需情况、产业结构四个方面综合选取了十二个指标，运用主成分分析法和多元回归研究了影响农产品冷链物流需求的若干因素，得到的模型方程拟合程度好，线性关系显著，能够较为精准的预测实际发生值，为山东省农产品冷链物流的发展提供一定参考。

关键词

山东省，冷链物流，主成分分析法，多元回归

Empirical Analysis of Cold Chain Logistics of Agricultural Products in Shandong Province Based on Principal Component Analysis and Multiple Regression

David Feng

School of Management, Guizhou University, Guiyang Guizhou

Received: Jul. 1st, 2023; accepted: Sep. 20th, 2023; published: Sep. 27th, 2023

Abstract

With the continuous development of economy, people's demand for safe and healthy agricultural

products is increasing. As a major agricultural province, Shandong Province has rich agricultural products resources, and the output of meat, fruits, vegetables and other agricultural products ranks among the top in the country. Taking the cold chain logistics of agricultural products in Shandong Province as the research object, this paper comprehensively selected 12 indicators from four aspects: regional economic development, transportation level, domestic and foreign market supply and demand, and industrial structure, and used principal component analysis and multiple regression to study several factors affecting the demand of cold chain logistics of agricultural products. The obtained model equation has a good fitting degree and a significant linear relationship. It can accurately predict the actual occurrence value and provide a certain reference for the development of cold chain logistics of agricultural products in Shandong Province.

Keywords

Shandong Province, Cold Chain Logistics, Principal Component Analysis, Multiple Regression

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着人们消费结构和模式的转变,发展农产品冷链物流,满足人们对高质量、多元化农产品的需求,是提高生活品质和幸福度的必然要求[1]。在山东省农产品供应的基础上,构建出影响冷链物流的指标体系并进行线性需求预测分析,有利于农产品质量安全的提升以及对冷链物流的发展具有一定引导作用。

① 冷链物流的影响因素。Can Eksoz (2018)将农产品产量、冷链流通率以及常驻人口数量等纳入了农产品冷链物流的影响因素中[2]。Elizabeth Gilliam (2020)认为影响农产品冷链物流发展最主要的因素是冷链物流的成本[3]。Catto Steven (2020)建议利用现有的数据建立模型,分析好每个变量之间的相互关系并赋予权重,根据实际不断改进、优化,以此反应冷链物流的运行效率[4]。Ya B (2016)运用系统顺序参数的多元回归和 AW-BP 预测,分析了食品冷链的影响因素以及需求[5]。文先明(2021)运用灰色关联法,用湖南省第一产业增加值、物流总费用、农产品市场成交量等因素分析了农产品冷链物流的需求[6]。

② 农产品冷链物流的评价指标。张晓茜(2021)将冷库容量、固定资产投资和农产品物流劳动力投入作为投入指标,分析农产品冷链物流业和货运量的增加值[7]。李倩倩(2020)将物流业固定资产投资、从业人员增长率以及人均冷藏库容量作为投入指标,分析了生鲜农产品冷链物流总额的增长弹性[8]。Jenkins Donald (2017)建立了包括顾客满意度、成本、坏损率、配送速度等 12 指标的评价体系,用于评估冷链物流环节的绩效[9]。

综上所述,现有研究主要集中在冷链物流效率评价体系,或构建各种模型分析冷链物流的影响因素上。但在现有研究中,对于山东省农产品冷链物流需求的研究不够。本文将在现有的研究基础上,收集共 12 个指标,运用主成分分析方法对影响指标进行降维,去除共线性,进而构建农产品冷链的预测模型,为山东省农产品冷链物流未来发展提供建议。

2. 变量及数据

2.1. 指标以及数据选取

① 区域经济发展水平。经济的发展水平直接影响着居民整体的消费能力和消费结构,决定了基础设施的建设程度和各个行业和产业的规模,人们对新鲜农产品品质的要求越高,就越能会促进农产品冷链

物流的发展。故选取以下指标：

X1: 地区生产总值(亿元)

X2: 第一产业增加值(亿元)

② 交通运输水平。交通运输在农产品冷链物流中起着关键性的作用，物流行业的运输能力和效率决定着冷链物流的发展，公路运输是农产品最为主要的流通方式，保证农产品的质量，需要实现全过程的温控，需要冷链设备、技术的支持。故选取以下指标：

X3: 货运量(万吨)

X4: 公路货运量(万吨)

X5: 货物周转量(亿吨公里)

X6: 公路运输业就业人员数(人)

③ 供需情况。供需情况影响了居民的消费习惯和消费意愿，居民的消费水平和可支配收入越高，对农产品的消费需求和品质要求也会随之升高，进而推动着农产品冷链物流的发展。故选取以下指标：

X7: 居民消费水平(元)

X8: 全体居民人均可支配收入(元)

X9: 社会消费品零售总额(亿元)

X10: 经营单位所在地进出口总额(百万美元)

④ 人文因素以及农业总产值。一般情况下，当地的人口数量很大程度上影响了农产品冷链物流技术的需求，常住人口越多，对农产品的需求也越多。农业的生产总值也同样影响当地产业结构，第一产业占比越大，农产品的供给和流通的需求就越多，反之亦然。故选取以下指标：

X11: 年末常住人口(万人)

X12: 农林牧渔业总产值(亿元)

指标具体定义如下表 1：

Table 1. Specific definitions of indicators

表 1. 指标具体定义

X1	指在一定时期内，该地区常住单位生产经营活动的成果。
X2	指在一定时期内，按市场价格计算的地区常住单位从事第一产业生产活动的成果。
X3	指在一定时期内，所有运输工具实际运送的货物重量。
X4	指在一定时期内，由各种公路运输工具实际运送到目的地并卸完的货物数量。
X5	指在一定时期内，由各种运输工具运送的货物数量与其相应运输距离的乘积之总和。
X6	指从事公路运输业并取得报酬或经营收入的人员。
X7	指按该地区常住人口平均计算的居民消费支出数值。
X8	指个人收入扣除相关税费以及交付政府的非商业性费用后的余额。
X9	指企业通过交易销售给个人、社会集团非生产、非经营用的实物商品金额，以及提供餐饮服务所取得的收入金额。
X10	指在所在地海关注册登记的有进出口经营权的企业实际进、出口额。
X11	指每年年末 12 月 31 日 24 时的人口数。
X12	指用价格体现的农、林、牧、渔业产品和对农林牧渔业生产活动的各种支持性服务活动的价值总量。

2.2. 数据查找

本文涉及到的数据均是通过国家统计局官网、山东省统计局官网以及山东省统计年报获取的，分别

是以上 12 个指标在 2003 年到 2021 年的年度数据, 具体数值如下表 2 所示。由于山东省冷链物流相关数据的统计较晚, 小部分数据缺失或不可获得, 考虑到数据的可获取性和可靠性, 部分缺失的数据由上一年数据与其同比增量的值相加而得, 以及通过总结现有的文献中的指标获取。

Table 2. Cold chain logistics demand index value of Shandong Province

表 2. 山东省冷链物流需求指标数值

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
2021	82875	6029	342728	291196	12049	201822	38547	35705	33715	453870	10170	11468
2020	72798	5364	317024	267230	10377	226659	36517	32886	29248	320210	10165	10190
2019	70541	5117	309533	266124	10166	203676	34627	31597	29251	296995	10106	9671
2018	66649	4951	354019	312807	10052	224934	30547	29205	27480	292397	10077	9397
2017	63012	4833	327006	288052	9719	225533	28353	26930	25528	264550	10033	9140
2016	58763	4830	285386	249752	8884	224205	25860	24685	23482	264550	9973	9075
2015	55289	4903	261849	227934	8418	224541	20684	22703	21551	240607	9866	9283
2014	50775	4663	264459	230018	8253	234715	19184	20864	25112	276929	9808	8988
2013	47344	4454	264100	227746	8194	234278	16728	19008	22295	266531	9746	8577
2012	42957	4047	333603	296754	11077	164478	15095	17127	19652	245544	9708	7817
2011	39065	3769	318407	279380	12684	107995	13524	15077	17156	235886	9665	7311
2010	33923	3411	301313	264366	11832	104576	11606	12922	14620	189156	9588	6573
2009	29541	3076	284086	251587	11022	101643	10494	11398	12363	139053	9470	5953
2008	27106	2876	244587	216604	10107	86503	9673	10411	10659	158407	9417	5583
2007	22718	2451	195259	163959	6413	79990	8142	9085	8608	122474	9367	4752
2006	18968	2098	164132	136750	6387	79899	7064	7795	7217	95213	9309	4058
2005	15948	1928	144701	120455	5551	83018	5916	6860	6167	76735	9248	3741
2004	13308	1748	129024	106887	4752	102767	4924	6355	4483	60658	9180	3453
2003	10903	1457	117051	97977	3908	103097	4351	5685	3937	44636	9125	2902

关于山东省冷链物流需求量, 在现有研究中, 有学者分别以生鲜农产品的产量与冷链流通率的乘积[10]或人均生鲜农产品消费量与人口数量的乘积[11]来代替。考虑到数据的代表性、有效性和可获取性, 本文选取了水果、肉类、水产品、蔬菜、牛奶的产量(万吨)总和来代替, 数据通过国家统计局中获取。具体数值如下表 3:

Table 3. Demand of cold chain logistics of agricultural products in Shandong Province

表 3. 山东省农产品冷链物流需求量

年份	2021	2020	2019	2018	2017	2016	2015	2014	2013	2012
Y	88090	85128	82127	80447	79308	77390	76176	74400	72321	70484
年份	2011	2010	2009	2008	2007	2006	2005	2004	2003	
Y	68609	65834	67403	66701	65352	61731	64172	62486	61297	

3. 对山东省农产品冷链物流的实证分析

3.1. 研究思路

农产品冷链物流一直是我国第一产业发展的重点, 对优化国民经济结构和改善居民生活具有重要作用。本文总结前人的研究, 运用 SPSS 软件, 将搜集到的 12 个指标通过主成分分析法降维为两个主成分,

将数据标准化后通过构建多元回归模型找到各主成分与冷链物流需求的相关关系，并利用模型并对其进行残差分析。

3.2. 主成分分析法

Table 4. KMO and Bartlett tests

表 4. KMO 和巴特利特检验

KMO 和巴特利特检验		
KMO 取样适切性量数。		0.825
巴特利特球形度检验	近似卡方	724.007
	自由度	66
	显著性	0.000

① 通过以上收集到的数据，运用 SPSS 软件进行主成分分析，如表 4 所示，KMO > 0.6，显著性水平 $p < 0.001$ ，说明指标之间存在相关关系，适合做主成分分析。

② 从下表 5 可以看出，本文共提取两个主成分，主成分 1 和主成分 2 的累积方差为 96.815%，表明能解释十二个指标信息的 96.815%，拟合优度高。

Table 5. Total variance interpretation table

表 5. 总方差解释表

成分	总方差解释					
	初始特征值			提取载荷平方和		
	总计	方差百分比	累积%	总计	方差百分比	累积%
1	10.491	87.427	87.427	10.491	87.427	87.427
2	1.127	9.388	96.815	1.127	9.388	96.815
3	0.224	1.867	98.682			
4	0.106	0.886	99.568			
5	0.029	0.239	99.807			
6	0.013	0.106	99.914			
7	0.006	0.049	99.962			
8	0.003	0.027	99.989			
9	0.001	0.005	99.995			
10	0.000	0.004	99.999			
11	0.000	0.001	100.000			
12	1.527E-5	0.000	100.000			

提取方法：主成分分析法。

③ 由于各指标的度量单位差异较大，需要将原始数据进行标准化处理，在此选择用凯撒正态化最大方差法旋转后的成分得分矩阵，并结合旋转后的主成分分析来解释因子，这里用 Y1、Y2 来代表主成分 1、2。旋转后的成分得分系数矩阵如下表 6 所示：

则可以得出：

$$Y1 = 0.274 * ZX1 + 0.264 * ZX2 + 0.158 * ZX3 + 0.149 * ZX4 + 0.063 * ZX5 + 0.289 * ZX6 + 0.276 * ZX7 + 0.280 * ZX8 + 0.271 * ZX9 + 0.248 * ZX10 + 0.268 * ZX11 + 0.262 * ZX12 \quad (\text{公式 1})$$

$$Y2 = 0.422 * ZX1 + 0.475 * ZX2 + 0.798 * ZX3 + 0.808 * ZX4 + 0.914 * ZX5 + 0.122 * ZX6 + 0.345 * ZX7 +$$

$$0.365 * ZX8 + 0.445 * ZX9 + 0.513 * ZX10 + 0.457 * ZX11 + 0.482 * ZX12 \quad (\text{公式 2})$$

其中, ZX_i 是变量 X_i 的标准化值。

Table 6. Component score coefficient matrix after rotation

表 6. 旋转后成分得分系数矩阵

	旋转后的成分矩阵	
	成分	
	1	2
X1	0.274	0.422
X2	0.264	0.475
X3	0.158	0.798
X4	0.149	0.808
X5	0.063	0.914
X6	0.289	0.122
X7	0.276	0.345
X8	0.280	0.365
X9	0.271	0.445
X10	0.248	0.513
X11	0.268	0.457
X12	0.262	0.482

3.3. 主成分多元线性回归分析

① 由上节得到 Y_1 、 Y_2 的得分, 带入山东省农产品冷链物流需求量 Y , 运用 SPSS 软件构建多元线性回归方程。

Table 7. Model coefficient table

表 7. 模型系数表

模型	系数 a					共线性统计	
	未标准化系数		标准化系数	t	显著性	容差	VIF
	B	标准错误	Beta				
(常量)	72077.083	422.129		170.747	0.000		
1 y3	4972.522	719.117	1.606	6.915	0.000	0.051	19.560
y4	-951.377	336.542	-0.657	-2.827	0.012	0.051	19.560

a. 因变量: y

由上述表 7 得知, 取显著性水平为 0.05, y_1 、 y_2 的回归系数 P 均小于显著性水平, 通过显著性检验, 回归模型可用。标准化的回归模型为:

$$Y = 4972.522y_1 - 951.377y_2 + 72077.083$$

结合公式 1、公式 2 可得:

$$Y = 963.23 * ZX1 + 862.38 * ZX2 + 28.36 * ZX3 - 27.55 * ZX4 - 554.73 * ZX5 + 1319.78 * ZX6 + 1045.94 * ZX7 + 1047.08 * ZX8 + 923.30 * ZX9 + 744.26 * ZX10 + 896.29 * ZX11 + 845.99 * ZX12 + 72077.083$$

② 为了使回归模型具有可比性, 将标准化的模型转化为非标准化的回归方程, 即:

$$Y_1 = 0.305 * X_1 + 0.306 * X_2 + 0.280 * X_3 + 0.274 * X_4 + 0.223 * X_5 + 0.260 * X_6 + 0.291 * X_7 + 0.299 * X_8$$

$$X8 + 0.306 * X9 + 0.299 * X10 + 0.305 * X11 + 0.306 * X12$$

$$Y2 = -0.130 * X1 - 0.070 * X2 + 0.381 * X3 + 0.405 * X4 + 0.641 * X5 - 0.401 * X6 - 0.197 * X7 - 0.187 * X8 - 0.106 * X9 - 0.008 * X10 - 0.090 * X11 - 0.060 * X12$$

$$Y = 0.036 * X1 + 0.032 * X2 + 0.003 * X3 + 0.001 * X4 - 0.017 * X5 + 0.048 * X6 + 0.038 * X7 + 0.039 * X8 + 0.034 * X9 + 0.028 * X10 + 0.033 * X11 + 0.032 * X12 + 53442.054$$

将 2003 到 2021 年的数据带入公式可以得出模型的预测结果，计算出实际值与预测值之间的残差，结果如下表 8。

Table 8. Model residual analysis

表 8. 模型残差分析

年份	实际值	预测值	残差
2021	88090.92	84921.82	3169.1
2020	85128.85	81508.11	3620.74
2019	82127.22	79511.79	2615.43
2018	80447.87	80136.42	311.45
2017	79308.32	78888.32	420
2016	77390.38	78247.67	-857.29
2015	76176.72	77035.89	-859.17
2014	74400.98	78366.70	-3965.72
2013	72321.61	77644.34	-5322.73
2012	70484.59	73575.49	-3090.9
2011	68609.72	70126.63	-1516.91
2010	65834.33	68128.56	-2294.23
2009	67403.02	66155.69	1247.33
2008	66701.06	65586.44	1114.62
2007	65352.04	63727.04	1625
2006	61731.06	62514.96	-783.9
2005	64172.14	61834.32	2337.82
2004	62486.29	62042.29	444
2003	61297.45	61391.89	-94.44

数据表明，预测值与实际值之间的残差较小，基本满足方差齐性和正态分布，经计算相对误差平均 2.5%，故在模型能较为正确的预测出实际的数值，模型精度高。

4. 结论以及建议

4.1. 结论

通过标准化后模型得知，经济发展因素、交通运输、供需情况、人文因素以及农业总产值均会对冷链物流的需求产生影响，除了 X4 公路货运量、X5 货物周转量即交通运输强度对山东省农产品冷链物流需求产生负面效益外，其他因素均产生正面影响。且从 2003 年到 2021 年的数据来看，各类农产品冷链物流的需求都在不断上升，属于冷链物流水平发展趋势较好的区域。

4.2. 建议

① 通过模型预测得知，未来的冷链物流需求还会不断增加，但现有的冷链基础设施、技术水平不足以满足未来需求，故政府应该加大资金投入，培养冷链物流专业人才，完善冷链基础配套设施，优化农

产品冷链物流系统。

② 鼓励企业发展并利用好冷链信息共享平台，共同促进山东省冷链物流信息管理系统的完善，拓宽山东省农产品的销售渠道，利用大数据提升冷链物流的运输效率，推动农产品产业结构的转型升级。

③ 利用山东省的区位优势，政府应该引导冷链物流与其他行业之间的合作，优化港口体系，把握地理上的优势，为冷链物流的发展提供新思路。

参考文献

- [1] 罗千峰, 张利庠. 农产品冷链物流高质量发展的理论阐释与实现路径[J]. 中国流通经济, 2021, 35(11): 3-11.
- [2] Eksoz, C., Mansouri, S.A., Bourlakis, M. and Önköl, D. (2018) Judgmental Adjustments through Supply Integration for Strategic Partnerships in Food Chains. *Omega*, **87**, 20-33. <https://doi.org/10.1016/j.omega.2018.11.007>
- [3] Gilliam, E.A. (2020) Optimization of the Cold Supply Chain Logistics Network with an Environmental Dimension. *IOP Conference Series: Earth and Environmental Science*, **628**, Article 012018. <https://doi.org/10.1088/1755-1315/628/1/012018>
- [4] Steven, C. (2020) A Novel Drug Quality Control Technology in Cold Chain Logistics Based on Port Transportation. *Coastal Education and Research Foundation*, **103**, No. 1. <https://doi.org/10.2112/SI103-142.1>
- [5] Ya, B. (2016) Study of Food Cold Chain Logistics Demand Forecast Based on Multiple Regression and AW-BP Forecasting Method on System Order Parameters. *Journal of Computational and Theoretical Nanoscience*, **13**, 4019-4024. <https://doi.org/10.1166/jctn.2016.4930>
- [6] 文先明, 肖锦. 基于灰色 GM(1, N)模型的湖南省农产品冷链物流需求预测[J]. 全国流通经济, 2021(6): 15-19.
- [7] 张晓茜. 京津冀农产品冷链物流效率与安全性研究[D]: [硕士学位论文]. 天津: 天津理工大学, 2021.
- [8] 李倩倩. 农产品冷链物流效率测度及提升策略研究[D]: [硕士学位论文]. 淄博: 山东理工大学, 2020.
- [9] Donald, J. (2017) Findings from Anhui Xinhua University Broaden Understanding of Chemistry (Research on the Mechanism of Cold Chain Logistics Subsidy). *Chemicals & Chemistry*, **79**, No. 3.
- [10] 柯亚楠. 唐山市农产品冷链物流需求分析与预测[D]: [硕士学位论文]. 武汉: 华中师范大学, 2013.
- [11] 王少然. 生鲜农产品冷链物流需求预测研究[D]: [硕士学位论文]. 西安: 西安工程大学, 2017.