

郑沫利,赵艳珂,陈思露,等. 基于 RDPSO-RBF 的粮食产后储藏环节损耗评估模型[J]. 江苏农业科学,2020,48(8):307-311.
doi:10.15889/j.issn.1002-1302.2020.08.056

基于 RDPSO-RBF 的粮食产后储藏环节损耗评估模型

郑沫利¹, 赵艳珂¹, 陈思露², 陈祺东², 孙俊²

(1. 国贸工程设计院, 北京 100037; 2. 江南大学物联网工程学院, 江苏无锡 214122)

摘要:为使减少粮食储藏环节的损耗,调查储藏环节中粮食作物的损失情况并进行评估预测。通过分析问卷,统计描述影响损失率的各个因素,应用随机漂移粒子群优化算法(random drift particle swarm optimization, RDPSO)和径向基函数(radial basis function, RBF)网络对储藏环节损耗进行评估预测,将调查问卷获得的数据作为 RBF 模型数据集,应用 RDPSO 算法进行模型的参数训练,从而获得粮食损耗的智能评估模型。采用稻谷数据作为各模型的数据集,通过仿真试验,对比数据拟合情况和均方误差,其 RDPSO-RBF 模型学习性能和泛化性能更强。因此, RDPSO-RBF 能更好地应用于实际的粮食管理中。

关键词:产后粮食;统计描述分析;RDPSO-RBF 模型;储藏损耗;评估

中图分类号: F326.11 **文献标志码:** A **文章编号:** 1002-1302(2020)08-0307-05

粮食是人类最基本的生存物资,是人类生活的第一要求,在国民经济中占有重要地位。长期以来,我国在保障粮食安全问题上一直较注重产前与产中的管理与投入。然而在粮食产后损失的关注严重不足,产后大概分为收获、运输、干燥、储藏、加工、销售和消费等环节^[1]。而每年粮食的产后损失情况极其严重,储藏环节粮食损失在产后总损失中所占比例较高。若在此环节能采取有力的防护措施,将可有效控制产后储藏环节中粮食的损耗。目前,已有不少的学者对粮食产后损失展开广泛的研究。在产后粮食损失方面,主要在储粮的损耗分析和减损措施。如储粮技术和储粮安全,主要是研究

不同因素对储粮损耗的影响,进而采取相关技术减损,保证储粮的质量与数量^[2-3]。在储粮模型构造方面,主要是依据机器学习或优化算法对储粮害虫进行研究^[4-6]。鲜有对粮食产后储藏环节的损失率预测的分析研究。

鉴于此,本研究调查了稻谷、大米、小麦、玉米、大豆 5 大粮食作物,通过问卷的形式记录了 10 个省份的产后粮食损失率和损失因素。为了解损失因素和其损失率的关系,将问卷进行变量处理和统计分析,并将各品种粮食作物生成变量相同的数据集。然后采用人工智能的方法,研究损失数据的影响因素对在储藏环节产后粮食损失率的影响,提出了基于 RDPSO(random drift particle swarm optimization)-RBF(radial basis function)损失预测模型。RBF 神经网络能很好地逼近非线性数据,真实地反映输入、输出变量间的内在联系,在众多神经网络中 RBF 逼近能力比较突出。但直接采用 RBF 作为损失率预测模型,其初始随机参数很难确定最优值,导致预测结果随机化。而 RDPSO 是一种全局优化算法,在搜索能力、收敛速度以及高维问题上表现突出。因此,用 RDPSO 优化 RBF 网络中参数,以之

收稿日期:2019-03-05

基金项目:国家公益性行业(粮食)科研专项(编号:201513004、201513004-6)。

作者简介:郑沫利(1967—),男,广东陆丰人,教授级高级工程师,主要从事粮食经济研究, E-mail: zhengmoli@sohu.com。共同第一作者:赵艳珂(1987—),女,河南洛阳人,博士研究生,中级工程师,主要从事粮食经济研究, E-mail: zhao_yanke@126.com。

通信作者:孙俊,博士,教授,主要从事人工智能、机器学习、计算智能和高性能计算研究。 E-mail: sunjin_wx@hotmail.com。

[24] Chapin P, Skalka C, Wang X S. Risk assessment in distributed authorization[C]//Proceedings of the 2005 ACM Workshop on Formal Methods in Security Engineering, 2005:33-41.

[25] Ni Q, Bertino E, Lobo J. Risk-based access control systems built on fuzzy inferences[C]//Proceedings of the 5th ACM Symposium on Information, Computer and Communication, 2010:250-260.

[26] Cheng P C, Rohatgi P, Wagner G M, et al. Fuzzy multi-level security: an experiment on quantified risk-adaptive access control[C]//IEEE Symposium on Security and Privacy, 2007:222-230.

[27] Hui W, Jing W, Shi B F. Model and application of green industry evaluation based on fuzzy control[J]. Journal of Intelligent & Fuzzy Systems, 2015, 29(6):2489-2494.

克服 RBF 网络预测结果的随机性,可使预测值更接近实际值。在实际储粮中,可以根据该预测值,判断是否需要采取措施来改善某些因素的影响,以达到减少粮食损耗的目的。

1 产后粮食储藏环节调查研究

1.1 产后粮食储藏环节损失率调查概况

以储藏环节的产后粮食作物稻谷、大米、小麦、玉米、大豆这 5 类品种作为调查主体。本研究根据全国粮食种植分布数据,抽取了青海、湖南、重庆、江西、福建、江苏、安徽、黑龙江、贵州、辽宁 10 个省份的数据,分别对这 5 类作物产后粮食的损失率情况进行专项调研。此外,在每个地区的抽样中选出几个市(区),对该抽样出的市(区)储粮企业单位进行访问。为了使问卷真实、有效,通过与企业代表人“一对一”访谈并当场答卷的方式进行调查,并由调查人员填写问卷。最后,调查得到有效的关于产后粮食在储藏环节损失率的问卷 324 份。其中,调查稻谷的样本问卷占总问卷的 36.7%,其次是玉

米、小麦分别占 29.3%、27.2%,最后大豆和大米占 4.3% 和 2.5%。因此,本研究选择稻谷作为代表粮食,分析影响其损失率的特征变量。

1.2 粮食产后储藏环节特征描述

在此次问卷中,该产后粮食损失率是对产后粮食总损失率的描述,即为产后粮食在储藏环节中的进仓损失率、出仓损失率、虫害损失率、鼠害损失率、霉变损失率、其他损失率的总和。而影响这些损失率的因素包含了受访者、受访企业的信息,产后粮食储藏的情况(年份、储藏品种、储藏时间、仓型、储藏量、储藏形式)、进仓损失信息(进仓工艺、进仓总量)、出仓损失(出仓工艺、出仓总量)等。从表 1 可以看出,粮食产后的储藏条件并不是很理想,其中储粮生态区域有 60.7% 为中温高湿储粮区;储粮技术有 59.4% 是尚未应用“四合一”技术的常规储藏;储藏形式采用散装的占 95.0%。同时,除了储藏仓型(使用平方仓的比例为 90.9%),其他特征所占比例分布相对较均匀。

表 1 特征的统计信息

特征描述	具体特征	频数	所占比例 (%)
企业所属储粮生态区域	高温高湿储粮区	22	10.0
	中温高湿储粮区	133	60.7
	中温低湿储粮区	6	2.7
	中温干燥储粮区	37	16.9
	低温高湿储粮区	21	9.6
储粮技术	尚未应用“四合一”技术的常规储藏	130	59.4
	应用“四合一”技术的储藏	84	38.4
	应用“四合一”升级新技术的储藏	5	2.2
年份	2014	66	30.1
	2015	62	28.3
	2016	89	40.6
	2017	2	0.9
储藏时间	1 ~ 20	83	37.9
	21 ~ 40	124	56.6
	41 ~ 60	12	5.5
仓型	露天囤	1	0.5
	地下仓	2	0.9
	钢板仓	5	2.3
	砖圆仓	4	1.8
	立筒仓	6	2.7
	浅圆仓	1	0.5
	平方仓	199	90.9
	其他	1	0.5

表 1(续)

特征描述	具体特征	频数	所占比例 (%)
储藏量(t)	1 ~ 1 000	45	20.5
	1 001 ~ 5 000	41	18.7
	5 001 ~ 10 000	46	21.0
	10 001 ~ 10 万	43	19.6
	10 万以上	44	20.1
储藏形式	散粮	208	95.0
	包粮	11	5.0
进仓工艺	散粮散装	157	
	散粮打包	3	
	包粮拆包	79	
	包粮包装	17	
进仓总量(t)	1 ~ 1 000	42	19.2
	1 001 ~ 5 000	47	21.5
	5 001 ~ 10 000	49	22.4
	10 001 ~ 10 万	44	20.1
	10 万以上	37	16.9
出仓工艺	散粮散装	202	
	散粮打包	7	
	包粮拆包	2	
	包粮包装	16	
出仓总量(t)	1 ~ 1 000	43	19.6
	1 001 ~ 5 000	47	21.5
	5 001 ~ 10 000	43	19.6
	10 001 ~ 10 万	44	20.1
	10 万以上	37	19.2

2 RDPSO - RBF 模型

2.1 RBF 神经网络简介

RBF 神经网络是一种前馈神经网络^[7]。一般最基本的 RBF 神经网络由输入层、隐含层、输出层构成 3 层网络结构,其公式如下:

$$Y=f(x)=\sum_{k=1}^K\alpha_k\varphi_k(x)。(1)$$

式中: α_k 是第 k 个隐含层节点到输出层的连接权值, $\varphi_k(x)$ 是第 k 个隐含层结点的输出值,其定义如下:

$$\varphi_k(x)=\exp\left(-\frac{\|x-\mu_k\|^2}{\sigma_k^2}\right)。(2)$$

式中: μ_k 和 σ_k^2 分别代表了第 k 个隐含层结点的中心向量和扩展常数。

2.2 RDPSO 算法简介

随机漂移粒子群(RDPSO)算法是基于 PSO 对

于受到自由电子运动轨迹分析的启发而提出的^[8-9]。当金属导体在外电场时,其导体内的自由电子会发生定向漂移,同时还存在无规则的随机热运动,这 2 种运动的结合可以使整体的势能逐渐变小。因此,电子的运动过程类似于优化问题中的求解最优解使得目标函数值最小化。在 RDPSO 算法中,第 $t+1$ 次的第 i 个粒子在 j 维的定向速度和随机热运动速度分别为 Vd_{ij}^{t+1} 、 Vr_{ij}^{t+1} ,因此第 $t+1$ 次粒子速度为:

$$V_{ij}^{t+1}=Vd_{ij}^{t+1}+Vr_{ij}^{t+1}。(3)$$

根据 PSO 算法中粒子趋向局部位置的学习机理,定向漂移速度 Vd_{ij}^{t+1} 的定义如下:

$$Vd_{ij}^{t+1}=c_1\cdot r_{ij}^t\cdot(P_{ij}^t-X_{ij}^t)+c_2\cdot R_{ij}^t\cdot(G_j^t-X_{ij}^t)。(4)$$

式中: c_1 和 c_2 为常数; r_{ij}^t 和 R_{ij}^t 是在 $(0,1)$ 区间内均匀分布的随机数, P_{ij}^t 是个体粒子最优值, G_j^t 是其所

有粒子中最优值。而式(4)可以等价

$$Vd_{i,j}^t = \beta \cdot (p_{i,j}^t - X_{i,j}^t); \quad (5)$$

$$\beta = c_1 \cdot r_{i,j}^t + c_2 \cdot R_{i,j}^t; \quad (6)$$

$$p_{i,j}^t = \frac{c_1 \cdot r_{i,j}^t \cdot P_{i,j}^t + c_2 \cdot R_{i,j}^t \cdot G_j^t}{c_1 \cdot r_{i,j}^t + c_2 \cdot R_{i,j}^t}。 \quad (7)$$

式中: β 为漂移系数; $p_{i,j}^t$ 为吸引子。

假设 RDPSO 中的粒子的无规则随机热运动速度 $Vr_{i,j}^{t+1}$ 服从双指数分布,通过蒙特卡洛法得到其表达式如下:

$$Vr_{i,j}^{t+1} = \delta_{i,j}^t \cdot \varphi_{i,j}^t。 \quad (8)$$

式中: $\varphi_{i,j}^t$ 为高斯分布函数; $\delta_{i,j}^t$ 是高斯分布的标准差,定义为

$$\delta_{i,j}^t = \alpha \cdot |M_{i,j}^t - X_{i,j}^t|。 \quad (9)$$

α 为热敏系数; $M_{i,j}^t$ 为当前种群中所有粒子个体的历史最优位置平均值,即

$$M_{i,j}^t = \frac{1}{N} \sum_{i=1}^N P_{i,j}^t。 \quad (10)$$

因此,式(8)可以改写为:

$$Vr^{t+1} = \alpha \cdot |M_{i,j}^t - X_{i,j}^t| \cdot \varphi_{i,j}^t。 \quad (11)$$

综上,RDPSO 算法中粒子速度和位置的更新公式为:

$$V_{i,j}^t = \beta \cdot (p_{i,j}^t - X_{i,j}^t) + \alpha |M_{i,j}^t| - X_{i,j}^t \cdot \varphi_{i,j}^t; \quad (12)$$

$$X_{i,j}^t = X_{i,j}^t + V_{i,j}^{t+1}。 \quad (13)$$

2.3 RDPSO - RBF 模型

2.3.1 RBF 神经网络模型确定 通过调查得到的问卷,在 SPSS 中进行分析。样本在变量处理后得到 51 个特征变量。由于某些粮食品种样本数量相对较少,特征数量过多,在试验中其误差结果并不理想。通过相关性分析,得到其中 8 个相关性最高的变量,作为实验室数据样本输入变量,而产后总损失率作为样本输出。因此,RBF 神经网络的输入层结点个数为 8,输出结点 1 个。隐含层通过利用对手受罚的竞争学习算法确定为 8。综上所述,该 RBF 神经网络模型结构为 8-8-1。

2.3.2 RDPSO 算法目标函数 采用均方误差作为 RDPSO 算法的目标函数,评价粒子群中所有个体,从中寻找最佳个体来判断是否更新粒子的 P_i^t 和 G^t 。

2.3.3 RDPSO - RBF 算法步骤 步骤 1:通过上述描述的思想,构建 RBF 网络,同时给定输入与输出样本集。

步骤 2:将 RBF 神经网络中的隐含层中心、扩展系数和权值编码成实数串代表粒子个体,同时随机产生规模为 S 的粒子群并初始化其位置和速度。

步骤 3:解码得到每个粒子个体串,映射到 RBF 网络的不同参数,同时将 N 组训练样本输入到 RBF 网络中,得到相应的输出,计算优化算法的目标函数。

步骤 4:计算漂移系数 β ,根据式(10)计算群体平均最优位置。

步骤 5:对每个粒子按照式(12)和(13)进行速度和位置的更新。

步骤 6:更新得到全局最优值 G_j^t 和粒子局部最优位置 P_j^t 。

步骤 7:判断全局最优值是否满足 RDPSO 算法中的结束条件,即算法给定预定值或迭代条件。如果满足条件,转入步骤 8,否则转入步骤 3。

步骤 8:算法结束。

3 结果与分析

基于上述思路在 MATLAB2015a 中编程,创建 RDPSO - RBF 网络模型,采用的计算机配置为: Intel® Core™ i7 - 6500U CPU @ 2.5 GHz, 8 G 内存。这里采用产后粮食作物稻谷作为试验数据集,将 207 组样本数据随机分成 2 份,前 160 组作为网络的训练数据,后 47 组作为测试样本。基于上述描述,通过降维后的变量(这里利用相关性,如:企业所属生态区域、储藏量、进仓总量、出仓总量、进仓形式是散粮散装进仓、出仓形式是包粮拆包出仓等)作为模型网络的输入,其产后粮食总损失率作为输出。

首先,通过训练样本训练出预测模型,并根据学习得到的模型输出的训练样本的预测值与训练样本实际值作比较。为了突出 RDPSO - RBF 模型预测产后总损失率的效果,将设计线性回归模型、RBF 模型进行试验效果的比较。从图 1 可以看出,线性回归模型在训练样本的拟合情况相当差,说明了该数据集并非线性的,线性模型在这里不适用于粮食的产后损失数据。而非线性的 RBF 神经网络和 RDPSO - RBF 网络,却能较好地拟合。

通过图 1 得知,RBF 神经网络算法和 RDPSO - RBF 算法通过训练样本学习到模型,其模型对训练样本有精准的预测输出。因此,为了检验模型的泛化性能,图 2 呈现了在测试集上的拟合情况,可以看出基于 RDPSO - RBF 模型能更好地预测损失率,其泛化性能比基于 RBF 神经网络模型的更好,其 RBF 模型有过拟合现象。

根据表 3 的 2 种算法模型的均方误差的数据可知 RDPSO - RBF 模型学习能力与泛化能力的优越

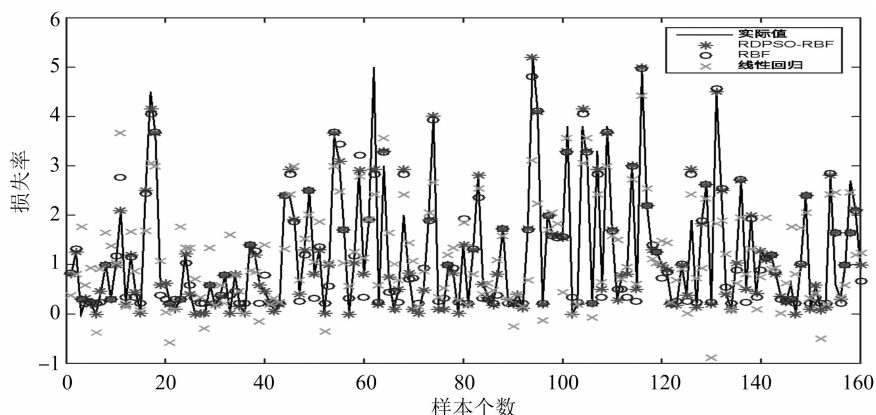


图1 训练样本拟合情况

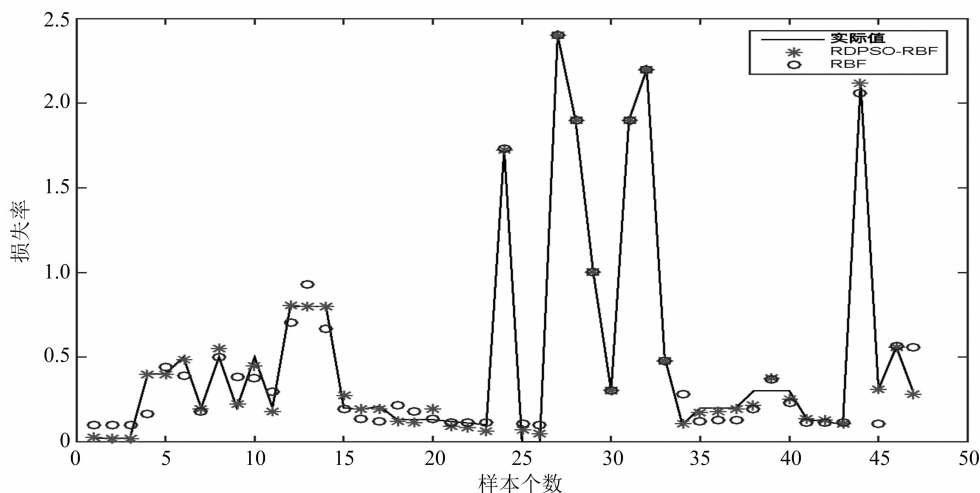


图2 测试样本拟合情况

表3 模型均方差比较

模型	训练均方误差(%)	测试均方误差(%)
RBF 神经网络	0.099	0.089
RDPSO-RBF 网络	0.049	0.019

性,能更好地预测粮食产后损失率。

4 结论

研究储藏环节粮食损失因素、预测粮食损失率,以期为粮食在储藏环节损失量提供参考值。为此,本研究基于对粮食产后损失率情况的调查,描述影响损失因素的特征。为精准地预测粮食产后储藏环节损失率,建立了 RDPSO-RBF 网络模型。通过试验对比,RDPSO 优化后的 RBF 神经网络能更好地对损失率作出准确预测,具有更强的泛化性能,在实际场景中预测损失率的值更加准确。

参考文献:

[1]高利伟,许世卫,李哲敏,等. 中国主要粮食作物产后损失特征及

减损潜力研究[J]. 农业工程学报,2016,32(23):1-11.

[2]张欢欢,唐浩林,叶福康,等. 粮食储存损耗的原因分析及对策[J]. 粮食加工,2015,40(3):69-72.

[3]张 彩. 散装粮食平房仓外围护体系研究[D]. 郑州:河南工业大学,2015.

[4]胡丽华,郭 敏,张景虎,等. 储粮害虫检测新技术及应用现状[J]. 农业工程学报,2007,23(11):286-290.

[5]Shen Y F,Zhou H L,Li J T,et al. Detection of stored-grain insects using deep learning[J]. Computers & Electronics in Agriculture, 2018,145:319-325.

[6]郭 敏,张明真. 基于 GMM 和聚类方法的储粮害虫声信号识别研究[J]. 南京农业大学学报,2012,35(6):44-48.

[7]Chen S,Cowan C F N,Grant P M,et al. Orthogonal least squares learning algorithm for radial basis function networks[J]. IEEE Transactions on Neural Networks,1991,2(2):302-309.

[8]Sun J, Wu X, Palade V, et al. Random drift particle swarm optimization algorithm; convergence analysis and parameter selection[J]. Machine Learning,2015,101(1/2/3):345-376.

[9]Kennedy J,Eberhart R. Particle swarm optimization:Proceedings of ICNN'95-International Conference on Neural Networks[C]. Perth: IEEE,2002,1942-1948.

[10]张 铃,张 钺. 人工神经网络理论及应用[M]. 杭州:浙江科学技术出版社,1997.