Multidimensional Classification for Systematization of Fish Processing Equipment

M A Saltykov,¹ T I Tkachenko²
¹Department «Process machines and equipment», Far East State Technical Fisheries University, Lugovaja, 52B, St., Vladivostok, 690087, Russian Federation
²Department «Process machines and equipment», Far East State Technical Fisheries University, Lugovaja, 52B, St., Vladivostok, 690087, Russian Federation

E-mail: saltykov_ma@mail.ru, tatkach_2002@mail.ru

Abstract. The paper dwells upon an approach to classifying fish processing machinery by means of cluster analysis. The classification presented herein helps optimize equipment upgrades and troubleshooting. The goal hereof is to analyze the fish processing machinery present in the market by means of cluster analysis. The paper covers fish dressing machines present in the fish processing market. The research methods include hierarchical cluster analysis, Euclidean distance, intergroup relations, and comparative analysis. SPSS is used to run calculations. The research has produced fish processing machinery clusters distinguished by performance, processable raw materials, and engine power; the research team has also tested a method for multidimensional classification for fish processing machinery analysis. These results are of use in a broad range of production tasks; they will be of interest for fish processing market analysts and machinery designers. Multidimensional classifications can be applied to arrange the production process, including selecting the most optimal equipment from a list of similar units.

1. Introduction

Russia’s economy is facing the need to upgrade its agricultural equipment, as the national agriculture lacks novel high-performance means of production. Agriculture is in dire need of upgrades, which should provide state-of-the-art machinery; as of today, only 20% to 30% of the machinery used in Russia’s agriculture can be considered state-of-the-art [1, 2]. Fishery is a subsector of agriculture; as such, it is facing the same problems: depreciation of production assets, lackluster performance, and low share of processed fish in the output. As an industry, fishery may opt for a technology-based approach to development, which implies increasing waste-free machine processing of fish [3].

To assess the existing processing machinery, its pros and cons, and to find out what technology could bring to the industry, one needs to analyze and classify the existing models of fish dressing units. This is why this paper attempts to apply multidimensional classification to fish processing machinery, including cluster analysis that splits a set of objects into homogeneous clusters. This method helps systematize and compare big data on the specifications of such machinery to better arrange and optimize the processing of aquatic organisms. Such calculations require a large database and extensive statistical computing. This can be done by advanced computational technology and software such as Statistica, SPSS, etc. Low-scale statistics can be calculated in Excel. Research presented herein effectively addressed two groups of objectives:
1. Simulate an optimal production process by varying technically similar equipment. Fish processing machinery is a rapidly growing market that is gaining ever more models and manufacturers offering a wide range of modifications that vary in specifications: Cabinplant A/S, Cretel, Grasselli, Gunther, Industrias Fac, Inwestpol, Karpowicz, Kittner, Maja, Marel, Nock, Rosoma, etc. The goal was to convert a list of machinery into a big data array that specifies a variety of technical parameters. Where different manufacturers offer a broad and diverse range of modifications, this might be a problem. Therefore, multidimensional classification based on information sheets for machinery analysis might be a solid solution.

2. The second objective was to analyze and classify fish processing machinery by the analyzed parameters such as performance, depreciation, engine power (kWh), power consumption, resource consumption (water, salt, chips, etc.), environmental parameters, size, weight, the number of operators, price per unit, and other parameters that could be optimized for when upgrading, troubleshooting, or choosing machinery for a facility.

2. Theory
This research used cluster analysis, a method presented in Bezdek J.C. [4], Dunn, J. [5], Everitt B. S. [6], Hartigan J.A. [7], Huang Z. [8], Jain A.K. [9, 10], Mirkin B. [11], Rand W. M. [12], Ward J. H. [13] et al.

Cluster analysis is common in engineering studies but not in food production research, which is why this paper draws upon engineering papers: Buchnev A. A. [14], Eltyshev D.K. [15], Kychkin A.V.[16], Khoroshev N.I. [17], Sivakov V. P.[18], Shtovba S. D.[19] et al.

3. Methodology
This research used hierarchical cluster analysis (hierarchical clustering), Euclidean distance, intergroup relations, and comparative analysis. It studied fish dressing machines. However, the methodology is applicable to many other machinery categories: smoking kilns, smoke generators, fish grills, skinning machines, injectors, heading machines, etc.

Fish machinery clustering can follow a two-step algorithm. Step 1 is to cluster by a single parameter: performance, or fish processing rate per minute. Step 2 is cluster by two parameters: performance and engine power. For classification, it should be borne in mind that each machine can run at minimum and maximum performance. The analysis thus can utilize the minimum value, the maximum value, and the mean. This research used the maximum value for clustering. Hierarchical clustering was performed in SPSS. The sample contained 18 most common fish dresser models; for some details on calculation methods used in fishery, see [20].

4. Practical significance
Clustering by performance has produced three clusters. Figure 1 shows a dendrogram that demonstrates machinery distribution by clusters as well as the distribution of singular performance values. Cluster 1: machines processing 200 to 320 fish specimens per minute. Cluster 2: 100 to 130 per minute. Cluster 3: 20 to 100 per minute. These clusters differ in terms of specifications and the type of raw materials the units could process, see Table 1 for the summary.
Cluster 1 (up to 320 fish specimens per minute) contains two machines from the sample: H2-ИРА110 and H2-ИРС. These units feature electric motors of up to 4.5 kW and weigh up to 1380 kg. They can process several types of fish. For instance, a H2-ИРА110 unit is designed to dress fresh and defrosted small fish varying from 140 to 260 mm in length; it can dress the sardine, the capelin, the mackerel, the navaga, the Baltic herring, the herring, the Pacific saury, the Atlantic horse mackerel, the hake, and some other species.

**Table 1. Fish dresser clusters.**

| Cluster | Unit | Performance, fish specimens per minute | Motor power, weight of the unit | Processable raw materials |
|---------|------|---------------------------------------|---------------------------------|--------------------------|
| 1       | Н2-ИРС, Н2-ИРА110 | 200-320 | 1.5 to 4.5 kW 1150 to 1380 kg | The capelin, the Baltic herring, the sardine, the Pacific saury, the sardinella, the mackerel, the saurel, the navaga |
| 2       | Н3-ИРОГ-2, Н3-ИРФ-2А | 100-130 | 0.5 to 3.55 kW 520 to 840 kg | The herring, the Alaska pollock, the capelin |
| 3       | RM-200, RM-350, ИПКС-074-01-140(Н), А8-ИФ4-Р, ИПКС-074-01-200-01(Н), КМ-50, H2-ИРА-125, FTC V-Cut, TAH-604, FGB-170, FGB-180, CHSF-1, CHCC-320 | 20-100 | 0.5 to 2.5 kW 70 to 400 kg | The herring, the cod, the haddock, the saithe, the hake, the Alaska pollock, the mackerel, the saurel, the trout, and salmons |

Cluster 2 comprises medium-performance machinery: three units that feature somewhat weaker engines, are lighter, and more specific, i.e. each unit being designed for one fish species. For instance, Н3-ИОГ-2 is designed to process herring material, Н3-ИРФ-2А is for the Alaska pollock, and H29-ИРШ is for the capelin.

Most of the models are in the third, low-weight, low-performance cluster. These units weight 150 to 250 kg on average. However, the range of processable fish is extensive here, although the capabilities vary from model to model. Some machines are species-specific. For instance, TAH-604 is
designed to dress the salmon, FGB-170, CHCC-320, and ИПКС-074-01-140(H) are for the herring, and FTC V-Cut is for the Salmonidae. Other models can handle multiple species: А8-ИФ4-Р is able to dress the cod, the Alaska pollock, the haddock, the trout, the saithe, the hake, the mackerel, and the saurel; H2-ИРА-125 can process the Gadidae, the Scombridae, the Sciaenidae, and the Clupeidae from 300 to 850 mm in length.

5. Conclusions
Thus, there is a fairly limited segment of high-performance machines, while the medium-performance models are species-specific. Undoubtedly, better clustering will require a larger sample and a greater range of parameters, which is the plan for further research.

Multidimensional classification based on the specification data of fish processing machinery is a potent tool that can identify both generic clusters and highly specialized clusters, as applying multiple classification metrics such as performance, raw materials, engine power, and resource consumption should be able to produce more specific clusters. The researchers plan to expand the model range and the analyzed parameters for further research. The results of these classification can be of use in:
- fish processing machinery database development and classification by a feature/features for design purposes;
- optimizing the choice of machinery for a specific process;
- addressing the environmental emissions by classifying the equipment by waste generation and applying specific anti-emission solutions;
- clustering the depreciated equipment at large-scale facilities to plan further upgrades and replacement.

6. References
[1] Burmistrova A and Rodionova N 2012 Problems of technical and technological modernization of agro-industrial complex in Russia Social-Economic Phenomena and Processes vol 1 (035) pp 33–37
[2] Polukhin A A 2015 Technical modernization of Russian agriculture in the conditions of international integration and economic sanctions Russian journal of agricultural and socio-economic sciences vol 6 (42) pp 41–51
[3] Tersky M V and Avdeev Yu A and Sergeev A S and Saltykov M A and others 2006 Competitiveness of the Russian Fisheries Complex Monograph (Vladivostok: Pacific strategy development center PSDC) 314 p
[4] Bezdek J C and Ehrlich R and Full W 1984 FCM: The fuzzy c-means clustering algorithm Computers & Geosciences vol 10 2-3 pp 191–203 doi.org/10.1016/0098-3004(84)90020-7
[5] Dunn J 1974 Well separated clusters and optimal fuzzy partitions Journal of Cybernetics vol 4 pp 95–104 doi:10.1080/01969727408546059.
[6] Everitt B S and Landau S and Leese M 2001 Cluster analysis (London: Arnold) 4th ed p 237
[7] Hartigan J A 1975 Clustering Algorithms (Wiley and Sons) p 351
[8] Huang Z 1998 Extensions to the k-means algorithm for clustering large data sets with categorical values Data Mining and Knowledge Discovery vol 2 pp 283–304 doi:10.1023/A:1009769707641.
[9] Jain A K and Dubes R C 1988 Algorithms for Clustering Data (Prentice Hall) p 320
[10] Jain A K and Murty M N and Flynn P J 1999 Data clustering: a review ACM Computing Surveys vol 31 3 pp 264–323
[11] Mirkin B 2011 Mathematical Classification and Clustering (Springer) p 448
[12] Rand W M 1971 Objective criteria for the evaluation of clustering methods Journal of the American Statistical Association American Statistical Association pp 846–850 doi:10.2307/2284239. JSTOR 2284239
[13] Ward J H 1963 Hierarchical grouping to optimize an objective function Journal of the American Statistical Association pp 236–244
[14] Buchnev A A and Pyatkin V P 2016 Hard and Fuzzy Clustering of the Earth Remote Sensing Data Journal of Siberian Federal University Engineering & Technologies Vol 9(7) pp 972-978

[15] Eltyshev D K and Gnutova K A and Litvinenko D A 2018 Automated evaluation of parameters for electrical equipment diagnostic models using fuzzy clustering Verm National Research Polytechnic University vol 28 pp 22–140

[16] Kychkin A V and Kvitko I I 2018 Architectural and technological organization of the information system for managing large data in industry and energy Vestnik Permskogo natsional’nogo issledovatel’skogo politekhnicheskogo universiteta. Elektrotekhnika, informatsionnye tehnologii, sistemy upravleniiia vol 25 pp 109–125.

[17] Khoroshev N I 2016 Intelligent decision support based on adaptive cluster analysis when exploiting power engineering equipment Sistemy. Metody. Tekhnologii Vol 3(31) pp 123–128 doi: 10.18324/2077-5415-2016-3-123-128

[18] Sivakov V P 2015 Cluster diagnostics during a comprehensive assessment of the operability of technical equipment Repair, restoration, modernization Remont. Vosstanovlenie. Modernizatsiya (Repair, Reconditioning, Modernization) vol 7 pp 17–19

[19] Shtovba S D and Pankevich O D and Nagorna A V 2015 Analyzing the Criteria for Fuzzy Classifier Learning Automatic Control and Computer Sciences Vol 49 3 pp 123–132 doi:10.3103/S0146411615030098

[20] Saltykov M A 2019 Guidelines for cluster analysis using information technology in fisheries research Naucnye trudy Dalrybvtuza vol 4 pp 103–110