

## Research on cognition and utilization mode of tail vegetable processing based on data mining technology

Ziying Tan, Shan Hu, Zhongshu Shi, Panan Zhang

School of Anhui University of Finance and Economics, Bengbu, China

### Abstract

The comprehensive treatment of organic matter is the key component of the deployment of ten key projects in the Notice of the Comprehensive Work Program of Energy Conservation and Emission Reduction in the 14th Five-Year Plan. Agricultural waste recycling represented by fruit and vegetable residues (i.e., "tail vegetables") has been paid more attention. In this context, we collected data through a combination of online and offline research, used data mining technology to explore the external guiding factors that affect the recycling efficiency, and text mining technology to focus on individual cases to explore suggestions on the recycling and utilization of tail vegetables from the perspective of residents. Finally, based on the data analysis conclusions and combined with the characteristics of different regions of China, suggestions were provided for the treatment of tail vegetables. Firstly, entropy method was used to construct an evaluation index system for the recovery efficiency of cabbage, and external guiding factors affecting the recovery efficiency in organizational governance were analyzed. Finally, text mining technology, LDA topic model and Apriori algorithm were used to explore the development prospects and difficulties of tail vegetable recycling from the perspective of residents. Based on the information of "characteristics of tail vegetables" and "residents' cognition", combined with the characteristics of tail vegetables in different regions of China, this survey proposed innovative tail vegetables treatment methods and promotion modes of deepening awareness, promoting source reduction, improving the organization supporting and mechanism, improving the organization system and regional planning, and deepening the awareness of environmental protection in the recycling process.

### Keywords

Tail vegetable recycling; Regional organization innovation; Grey correlation analysis; Entropy method.

### 1. Introduction

"The Ministry of Agriculture and Rural Affairs on the implementation of the Central Committee of The State Council in 2023 to comprehensively promote the deployment of the key work of rural revitalization" requires the implementation of rural sewage, waste agricultural film recycling and disposal of vegetables, village planning, "four good rural road" construction, village public welfare facilities co-management and sharing, livestock and poultry breeding waste and straw resources utilization, and other "six actions." In November 2021, The State Council issued the 14th Five-Year Plan for Modernization of Agriculture and Rural Areas, which lays out a comprehensive arrangement for the strategic orientation, main objectives, key tasks, policies and measures for modernization of agriculture and rural areas during the period. The "Plan" proposes to promote green production mode, promote the reduction of inputs, clean production, waste resources, ecological production mode, and fundamentally solve the problem of agricultural non-point source pollution. At the same time, advanced technology, modern equipment and management concepts will be introduced into agriculture,

infrastructure and basic public services will be extended to rural areas, agricultural production efficiency will be improved, rural appearance will be improved, farmers' quality of life will be enhanced, comprehensive agricultural upgrading will be promoted, the whole agricultural industry chain will be developed, the value chain will be expanded, comprehensive benefits will be enhanced, and sustainable agricultural development will be promoted.

From the perspective of agricultural environment, waste water, odor and bacterial spread caused by non-standard treatment of tail vegetables will cause agricultural non-point source pollution. From the perspective of agricultural governance, tail vegetable treatment is not only related to environmental security and resource reuse, but also closely related to the sustainable development of fruit and vegetable industry and rural revitalization strategy. From the perspective of the development of agricultural technology, the agricultural machinery and equipment represented by tail vegetable treatment and the updating and iteration of processing and utilization technology will promote the improvement of agricultural industrial chain and the expansion of value chain. Fundamentally improving the recycling function and utilization efficiency of tail vegetables is an important foundation for promoting agricultural and rural modernization, and a key measure for building a beautiful China, which will help accelerate the improvement of agricultural modernization.

As early as 2018, the No. 1 document of the Central Committee clearly required the implementation of the three-year action plan for the improvement of rural living environment, focusing on the treatment of rural garbage, sewage and the improvement of village appearance, integrating various resources, strengthening various measures, and steadily and orderly promoting the treatment of prominent problems in rural living environment.

In the development and utilization of resources, the major themes of resource conservation, prevention and treatment at the source, and ecological restoration echo the targets of source reduction, environmental recovery and environmental optimization of utilization and management of tail vegetable. The construction of recovery and treatment of tail vegetable is promoted to conform to the direction of national ecological civilization construction and walk at the forefront of environmental improvement.

As one of the components of rural garbage and sewage treatment, the optimization of tail vegetable treatment is of great practical significance to improve the quality of living environment and promote the construction of ecological environment. Li Yanming, associate professor of China Agricultural University, has pointed out that it is urgent to turn tail vegetable into treasure and make effective use of it in view of the rotting of tail vegetable in some vegetable producing areas and rural areas. Improving rural living environment is an important part of rural revitalization and ecological civilization construction.

## **2. Construction of evaluation index system of tail vegetable recovery efficiency based on entropy method**

### **2.1. Research Ideas**

Based on the evaluation index system of tail vegetable recovery efficiency, since the units of measurement of each index are not uniform, we first processed the data in a dimensionless way. Then the entropy method was used to calculate the weight of each index and reflect the influence of each index on the recovery efficiency of cabbage. Then, a comprehensive evaluation model was established to calculate the comprehensive score of tail vegetable recovery from 2016 to 2020. Finally, according to the weights of each index and the scores of each year, the current situation and existing problems of tail vegetable recycling were analyzed based on relevant literature and data and previous research results.

### 2.2. Construction of recovery efficiency evaluation index system

Do not number your paper: All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper. Use italic for emphasizing a word or phrase. Do not use boldface typing or capital letters except for section headings (cf. remarks on section headings, below).

Table 1: evaluation index system of recovery efficiency of tailing vegetable

First-order index	Secondary index	symbol	unit
Propaganda intensity	Frequency of publicity and training	$x_1$	time
	Number of publicity and training	$x_2$	persons
	Vegetable planting area	$x_3$	Ten thousand hectares
Planting scale	Vegetable yield	$x_4$	Ten thousand tons
	Yield of tail vegetable	$x_5$	Ten thousand tons
	Tail vegetable treatment area	$x_6$	Ten thousand hectares
Governance situation	Utilization of tail vegetable treatment	$x_7$	Ten thousand tons
	Utilization rate of tail vegetable processing	$x_8$	%

### 2.3. Establishment of the model

In this paper, the data used includes both positive and negative indicators. In order to prevent 0 between positive and negative indicators, we need to improve the normalization formula of the general evaluation model.

For positive indicators:

$$\lambda_{ij} = 0.98 \frac{\lambda_{ij} - \min(\lambda_{ij})}{\max(\lambda_{ij}) - \min(\lambda_{ij})} + 0.02$$

For negative indicators:

$$\lambda_{ij} = 0.98 \frac{\max \lambda_{ij}}{\max \lambda_{ij} - \min \lambda_{ij}} + 0.02$$

Then calculate the proportion of the i sample in the j index:

$$p_{ij} = \frac{\lambda_{ij}}{\sum_{i=1}^5 \lambda_{ij}}$$

Then the information entropy of each index can be written as:

$$e_j = -\frac{\sum_{i=1}^5 p_{ij} \ln p_{ij}}{\ln 8}$$

The weight of each indicator is:

$$w_{ij} = \frac{1 - e_i}{n - \sum e_i}$$

Through calculation, we get the weight of influencing factors of recovery efficiency as follows:

Table 2: Weight of factors affecting recovery efficiency of tail vegetable

index	Entropy value	Coefficient of difference	weight
Frequency of publicity training (times)	0.8227	0.1773	0.1051
Number of publicity and training personnel (persons)	0.7834	0.2166	0.1284
Vegetable planting area (ten thousand hectares)	0.6738	0.3262	0.1934
Vegetable production (10,000 tons)	0.7354	0.2642	0.1568
Tail vegetable production (10,000 tons)	0.8573	0.1427	0.0846
Area under control of tail vegetables (ten thousand hectares)	0.8580	0.1420	0.0842
Utilization of tail vegetable treatment (10,000 tons)	0.8476	0.1524	0.0903
Utilization rate of tail vegetable treatment (%)	0.7346	0.2654	0.1573

The total index and the three first-level index scores in each year were calculated respectively:

$$F = \sum_{j=1}^8 w_j x_{ij}$$

Table 3: Weight of factors affecting recovery efficiency of tail vegetable

annual	F	F <sub>1</sub>	F <sub>2</sub>	F <sub>3</sub>
2016	0.2445	1.0000	0.0238	0.0020
2017	0.5818	0.6772	0.4913	0.6334
2018	0.6263	0.4498	0.4359	1.0000
2019	0.4350	0.3160	0.5519	0.3655
2020	0.6795	0.0020	0.9703	0.7750

### 2.4. Multi-dimensional analysis of the evaluation results of the recovery efficiency of cabbage

As can be seen from the above table, the publicity and training of tail vegetables show a downward trend from 2016 to 2020, which is because the early training is in place enough to

make people have a deeper awareness of tail vegetables, not only treat it as garbage, understand the value of tail vegetables, and consciously reduce the waste of tail vegetables in daily life. Improve the recovery efficiency of tail vegetable field.

The current way of publicity is mainly technical training. Take Gangu County as an example, on March 25, Gangu County agricultural technology promotion center seize the opportunity, convened Panan town Xie Jiaping, ten Jiaping, Pei Jiaping, Li Jiaping village vegetable farmers, vegetable growers and planting professional cooperative personnel, held tail vegetable processing and use of project technical training. At the training site, professional technicians from the county agricultural Technology promotion center explained the compost technology, compost technology for compost vegetables, composting technology for compost vegetables and composting technology for compost vegetables directly back to the field as well as precautions, and demonstrated the relevant technical operation processes such as compost treatment for compost vegetables and compost utilization. Through compost and compost fermentation of compost vegetables and compost vegetables, non-point agricultural source pollution was reduced, soil organic matter was increased, and sustainable circular agriculture was built. Take this opportunity, professional and technical personnel to farmers issued waste agricultural film recycling regulations, propaganda waste agricultural film recycling policy and the importance of agricultural film recycling.

However, the publicity methods of tail vegetable processing and recycling are relatively simple, mainly technical training, and the publicity and training in the later stage show a declining trend. We should timely improve this situation, make full use of various media such as the Internet, and increase publicity efforts.

In terms of planting scale, due to the continuous adjustment of agricultural industrial structure and the continuous improvement of people's living standards, vegetable planting area, yield and tail vegetable output also gradually increased. The increase of the output of tail vegetables also increases the difficulty of recycling tail vegetables.

On the management of tail vegetables, the management of tail vegetables can be roughly divided into two stages. From 2016 to 2018, the management of tail vegetables showed a steady rising trend. The treatment area, treatment utilization and utilization rate of tail vegetable showed an increasing trend. At present, the suppressant treatment methods mainly include compost method, compost method, and compost method.

### **3. Explore the development prospects and difficulties of tail vegetable recycling from the perspective of residents based on data mining technology**

#### **3.1. Innovative exploration of tail vegetable recycling based on text mining**

We use TF-IDF algorithm to deeply understand the areas that still need to be improved in tail vegetable recycling from the perspective of residents. After jieba segmentation and jieba removal by Python, key words were extracted, and the results showed that the residents supported the suggestion of vegetable recycling, while some residents had a negative attitude, and some words appeared such as "nothing" and "nothing to do with me". However, most residents are supportive and actively offer suggestions for the recycling of tail vegetables. Words such as "publicity", "regionalization" and "classification" appear.





Figure 2:LDA topic model solving

Through the analysis of the above results, we can conclude that most residents actively support the recycling of tail vegetables, and actively give suggestions for the disposal of tail vegetables, and put forward corresponding suggestions. A large number of residents believe that the lack of relevant waste disposal facilities and related technologies still need to be improved is one of the important problems facing tail vegetable recycling at present. At the same time, residents generally believe that the treatment of tail vegetables can effectively improve the environment, and actively respond to the national call for energy conservation and emission reduction. The classification and recycling of tail vegetables can realize the secondary utilization of resources, but some systems and supervision mechanisms still need to be further improved.

### 3.3. Difficulty analysis of tail vegetable recovery based on Apriori association rule

In order to have an in-depth understanding of the biggest problems faced by tail vegetable recycling from the perspective of residents, we used Apriori algorithm to analyze the data obtained.

Item set: A nonempty collection of data items in the database.

Transaction: A transaction contains one or more item sets.

Support: Number of transactions containing item set  $x$  as a percentage of total transactions.

Confidence: The ratio of the number of transactions containing both data item  $x$  and data item  $y$  to the number of occurrences of transaction  $x$  (or transaction  $y$ ).

Minimum support and minimum confidence: Minimum requirements that association rules must meet.

Frequent item set: The item set that is greater than or equal to the minimum support is called frequent item set.

Maximum frequent item set: Frequent item sets that are not included by other frequent item sets are called maximum frequent item sets.

Strong association rules: greater than or equal to the minimum support and minimum confidence is called strong association rules (frequent item set based on greater than or equal to the minimum confidence).

For the convenience of subsequent calculation, a is denoting "weak public awareness", b is denoting "insufficient supporting facilities for relevant waste treatment", d is denoting "lack of advanced treatment technology" and d is denoting "lack of management and supervision mechanism".

Using Python, we get the following result:

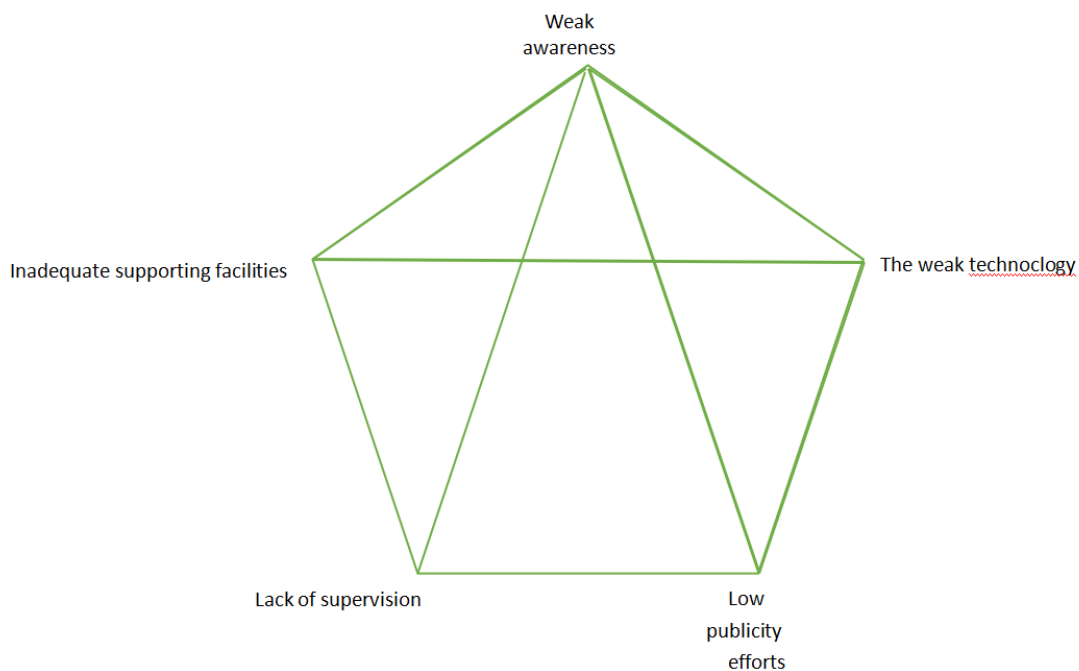


Figure 3: Apriori algorithm result graph

Through the results, we found that the primary problem at present is the weak public awareness. As a result, residents have little understanding of the recycling methods of tail vegetables. As a result, residents have weak awareness and do not deliberately separate tail vegetables from other garbage. Similarly, residents have no awareness of special treatment of tail vegetables, so the relevant knowledge of tail vegetables treatment can only be spread through some TV media. At the same time, the weak public awareness and the lack of management mechanism also have a certain connection. If residents are not aware of special treatment of tail vegetables, there will be no supervision in daily life. At the same time, the related waste disposal facilities are not in place, the treatment technology is not advanced enough is also facing the difficulties of tail vegetable recycling.

## 4. Suggest and summary

### 4.1. Research conclusion

This survey focuses on farmers, distributors and consumers related to the vegetable and fruit industry. Through descriptive analysis of the survey and questionnaire, combined with the current background of The Times, it is understood that the resource utilization of tail vegetables is facing the following similar situations.

#### 1. Residents have little knowledge of tail vegetable treatment

In the process of investigation, we learned that the public's understanding of the resource utilization of tail vegetables is generally not high, especially consumers. This reflects that Chinese residents pay little attention to the treatment of tail vegetables and the country's



publicity for the utilization of tail vegetables resources is not enough. Consumers are one of the main sources of scattered tail vegetables. If consumers have little environmental awareness and do not understand the relevant knowledge of tail vegetables treatment, they will add a lot of troubles and workload to urban waste treatment, reduce the possibility of effective reuse of tail vegetables, and have a negative impact on the environment. Only by improving consumers' awareness of environmental protection and cultivating their ability to deal with tail vegetables, can we better solve the problem of dealing with a large number of tail vegetables in the consumption link.

#### 2. The construction of supporting facilities for tail vegetable treatment is backward

In the process of investigation, it was found that people generally reflected that the supporting infrastructure construction of tail vegetable processing is backward, and the classification and recycling is not convenient. Due to the backward construction of tail vegetable treatment infrastructure, residents who contact tail vegetable treatment will not pay attention to tail vegetable treatment, which brings great hidden dangers for tail vegetable treatment. At present, urban and rural tail vegetable treatment facilities are poor, the management system is not perfect, is lack of supervision system, tail vegetable treatment is still a big problem. It is necessary to strengthen the infrastructure of waste treatment and the publicity to the public for the treatment of tail vegetables and the protection of the environment.

#### 3. Garbage disposal capacity is weaker in some parts of our country

Because the production of large quantities of tail vegetables is concentrated in the fruit and vegetable planting areas in western China, the western region has certain advantages for the concentrated treatment of large quantities of tail vegetables. However, as the production of tail vegetables involves various production and sales links, covering a wide range of scattered, so only some areas of centralized processing capacity is far from enough. At present, many citizens are not satisfied with our tail vegetable processing ability, especially in the retail link. This reflects that the treatment of tail vegetables in communities is not in place, the government does not pay enough attention to the problem of tail vegetables treatment, and the relevant technology is not advanced enough. Therefore, how to really deal with the end dish produced in each link, so as to satisfy the people, is the need for national and local departments to make concerted efforts and joint efforts.

### 4.2. Put forward suggestions

Combined with the characteristics of treatment technology and the influence of various factors such as agricultural resource endowment, geographical environment and climatic conditions in the north and south, the agricultural waste or residual resource utilization represented by tailing vegetable can be promoted and applied through different regional division, so as to improve the application of technology according to local conditions and enhance the treatment efficiency of tailing vegetable. According to terrain, rainfall, soil, cultural environment and other factors, agricultural waste can be divided into four categories, and combined with different kinds of basic technologies such as biogas, fertilizer, feed and other application adaptability analysis:

At present, agricultural waste resources processing and utilization mode is single, low efficiency, according to the terrain climate and other planting and breeding conditions, major agricultural waste division can be carried out in various regions, and targeted resources treatment technology and suggestions.

#### Division I: Southern hilly region

Planting industry and tail vegetable production status: Hilly and mountainous terrain reduces the possibility of large-scale grain planting, and the development of livestock and poultry is dominated by small family-oriented units. Cash crop planting represented by tea orchards

planted on gentle slopes and part of flat land becomes the main direction of agricultural development, while tail vegetable production is relatively small.

Treatment technology application suggestions: Under the condition that the yield of tail vegetable is not abundant, tail vegetable treatment in southern hilly areas can be combined with the common one-household operation in the local area. At the level of fertilizer treatment, it can be considered to mix with straw and other dry agricultural residues to achieve the neutral content of nutrients and water, and use livestock and poultry manure to increase fertilizer nutrients, so as to achieve a reasonable ratio and harmless local composting. Make up for the lack of fertility caused by soil resources and soil erosion; In the process of feed chemical treatment, due to the traditional one-family production economy in the southern hilly area, household digestion of tail vegetables has become the main way, and the nearby collection is promoted, and the straw of corn and wheat is mixed to improve the utilization rate and applicability of feed. Due to the warm and humid climate, tail vegetables will produce acidification characteristics in the humid environment of high climate. Anti-acidification treatment is helpful to keep the internal nutrients and quality of tail vegetables from being damaged and optimize the treatment efficiency and satisfaction of tail vegetables.

#### Division 2: Southern Plain area

Planting industry and tail vegetable production situation: Due to the relatively flat terrain, the planting area is more concentrated in the southern plain area. However, due to the dense water network, large-scale organization and processing are inconvenient, and the domestic economy in the planting level is slightly weaker than that in the hilly area. In terms of crop yield, high-moisture crops represented by fruits and vegetables account for the same proportion as other food crops due to environmental adaptability. Taking Anhui Province as an example, it has a multi-plain terrain in the north of Anhui Province. In 2021, vegetable and fruit output in Anhui Province accounts for 43% of all crop types and 90% of total cash crop output. Most of the planted products were products with rich water content, and the high proportion of cash crop yield was associated with the increase of high humidity of tail vegetable yield.

Suggestions on application of treatment technology: Since the distribution of cultivated land in the southern plain area is relatively dispersed but the yield of tail vegetable is high, it is difficult for farmers to treat and utilize tail vegetable depending on their individual strength, which is likely to lead to tail vegetable rot and non-point source pollution caused by sewage. Therefore, it is more necessary for regional organization collection. Tail cabbage is rich in yield and has typical characteristics of high water content and high organic matter. In the application of fertilizer, attention should be paid to dehydration treatment, and the mixed treatment mode in hilly area should be referred to, including straw and manure contamination for nutrient enrichment and water dilution. On the basis of mixed dehydration, the grinding process can be added in the application of feed to make full use of tail vegetable resources to produce easily stored feed. Under the condition of centralized agriculture in some areas, biogas can be utilized through small and medium-sized biogas fermentation mechanism, and after harmless treatment by aerobic fermentation equipment, biogas residue is used as cultivation substrate to make up for cultivated land and soil resources, and biogas slurry is used as organic fertilizer fertilization equipment in operation. At the same time, due to the characteristics of low soil water demand and weak digestion capacity of biogas slurry in humid environment and plain water network area, attention should be paid to the integration of efficient treatment of livestock and poultry wastewater and standard discharge technology in the application of feed and energy, so as to reduce agricultural non-point source pollution.

#### Division three: northern latitudes

Planting industry and tail vegetable production status: The northern mid-latitude region is partly similar to the southern plain region in terms of terrain and topography, but the terrain is more open and the water network is sparsely divided. Professional farming is more popular,

but dry land farming is the main agriculture, mainly producing wheat and corn and other food crops. In some areas, such as Shandong, vegetable production is concentrated, and tail vegetable production is relatively large. Concentrated tail production also puts pressure on the regional environment. Due to the restriction of low precipitation on soil water content, the main direction of treatment and utilization of tail cabbage is to make full use of the water resources of tail cabbage returning to the field and enhance the fertility and water content of organic matter combined with relatively developed animal husbandry.

Treatment technology application suggestions: In the preliminary collection process, large-scale farming in northern China requires the same large-scale organization and recycling to avoid seasonal tail vegetable pollution. Due to the transitional climate characteristics of the mid-latitude region in northern China, the corruption in the process of feed processing may be higher than that in the northern high-cold region, and the time concentration of tail vegetable production. It is helpful to save nutrient loss and prolong the storage time by mixing tailing vegetable and straw in anaerobic fermentation. In the process of fertilizer and energy, in view of the overall lack of water in the environment, the use of wet single-phase anaerobic digestion technology and equipment with large output of biogas slurry, the biogas slurry as a liquid organic fertilizer is applied to the nearby through large and medium-sized fertilization equipment, or with the integration of water and fertilizer equipment applied to the cultivated land, the use of water content resources of tail vegetables, enhance soil moisture and fertility, Biogas residue can be further processed by aerobic fermentation equipment to produce high-end organic fertilizer.

Division four: northern alpine region

Planting and tail vegetable production: The three eastern provinces have the largest agricultural scale and the strongest professional facilities. Due to the project of transporting western vegetables to the east in Inner Mongolia and Gansu, vegetables and potatoes are the main agricultural products in Gansu. Due to the wide terrain and high scale planting degree, tail vegetable output is higher than other regions, and collection and processing will also face more obstacles. Due to the temperate continental climate, low soil fertility, low water content, and more developed animal husbandry, the comprehensive utilization of tail vegetables and animal husbandry waste resources has become the main direction of regional development.

Application suggestions of treatment technology: As the main products of the northwest plateau region, vegetables and other main agricultural products, the production of common tail vegetables also requires higher scale collection procedures. Because the scale of animal husbandry in the northwest alpine region is more significant, waste resources of animal husbandry can also be collected in bulk. In the pretreatment process of tail vegetable, the dry continuous anaerobic digestion technology with high heat utilization rate and less biogas slurry production was used to achieve water and energy savings. In the process of fertilizer, tail vegetables, fecal pollution and wastewater are involved in the treatment of tail vegetables. Due to the cold and dry climate, the biogas process is carried out by aerobic fermentation and the use of biogas residue and the remaining tail vegetables for fertilization.

## References

- [1] Huang B, NIU M F. Research progress of resource utilization of tail vegetable [J]. Agricultural Technology and Equipment, 2021(08):70-71. (in Chinese)
- [2] Wang Mengzhi, Yang Nan, Yan Lu, Ding Ruoyan. Tail vegetable food feed the use and its impact on the animal body [J]. The research progress of feed industry, 2023, 44 (05) : 1-7. DOI: 10.13302 / j.carol carroll nki fi. 2023.05.001.

- [3] CChang Fanghong. In dingxi city tail food treatment and utilization present situation and development [J]. Journal of agricultural science and technology and information technology, 2023 (01) : 95-98. The DOI: 10.15979 / j.carol carroll nki cn62-1057 / s. 2023.01.033.
- [4] Wu Feng-xia, WANG Xiao-Xue, Xiao Qiang, ZHANG Shu-bin, LIU Jian-bin. Effects of different feeding methods on composting effect of cabbage [J]. Journal of Shandong Agricultural University (Natural Science Edition),2022,53(03):445-453.