ISSN 2618-1584 Vol. 3, Issue 4: 1-5, DOI: 10.25236/FMSR.2021.030401

# Research on the Model of Fungus Decomposition Rate of Wood Fiber Based on Principal Component Analysis Algorithm

**Ziqian Yang, Wenlu Xie, Shibao Chen**

*Guizhou Medical University, Guiyang, Guizhou 550000, China*

***Abstract:*** *Regarding the lignin fiber decomposition rate under the interaction of multiple fungi in the short term, was simplified as the interaction of two bacteria without the influence of external regulation, based on the two shapes of growth rate and moisture tolerance. Based on the growth rate data of fungal pair interaction, the decomposition rate model of lignin fiber based on principal component analysis was established, and the contribution ranking of fungi to decomposition rate was obtained. Then the moisture tolerance data were fitted to find the difference between the water niche width and the competition ranking of fungi. Based on the ranking of the two traits, a linear weighting equation was established, and the cumulative decomposition rate of the interaction between various bacteria was finally obtained to be1298.7.*

***Keywords:*** *principal component analysis, Decomposition rate of fungi, Bacteria ranking weight*

## 1. Introduction

Carbon cycle is an important part of life on the earth, which plays a crucial role in the stability of the ecosystem. The key part of carbon cycle is the decomposition of plant materials and wood materials, and the key factor in the decomposition of wood fiber is fungi. [1] Some characteristics of fungi determine their decomposition rate, such as the growth rate and humidity tolerance of fungi. [2] These two characteristics are closely related to the decomposition rate. In general, there is a linear relationship between them. And some traits also have an impact on the relationship, such as slow growth of fungi can better adapt to the changes in the environment, while the faster growth of fungi is on the contrary. [3]

## 2. Model establishment and solution

The decomposition rate of lignocellulose is mainly determined by the internal conditions (growth rate, humidity resistance) and external conditions (temperature and humidity). It is assumed that in the short term, the decomposition rate of wood fiber can be ignored by external conditions, and only the growth rate and moisture resistance are considered in the final decomposition rate. In order to simplify the model, we only consider the pairwise interaction between fungi.

- **Growth rate**

The relationship between growth rate and decomposition rate of 37 kinds of bacteria is known, f1, ..., f37 respectively represent 37 kinds of bacteria. Transforming the data in the data table into a matrix can more intuitively see the growth rate after the interaction between the two bacteria groups. The transformation of the data table is based on the pychar platform. Each value in the matrix is the growth rate corresponding to the interaction of the two bacteria. The original data is a pairwise comparison of 37 * 34, which is converted to 37 * 37. The combinations without mutual comparison are filled with 0.

Principal component analysis is to use the idea of dimension reduction to get the most representative index (that is, principal component) after the effect of bacteria. Here, the purpose of principal component analysis is to get a ranking of the weight of growth rate of various bacteria. Assuming that a1, ..., an represents the weight of the extracted principal component, the comprehensive score of each bacterium can be expressed as

$$score = a_1 \times x_1 + a_2 \times x_2 + \cdots a_n \times x_n$$

When one principal component is not enough to represent the original variable, the second or even

ISSN 2618-1584 Vol. 3, Issue 4: 1-5, DOI: 10.25236/FMSR.2021.030401

more principal component information will be searched iteratively. The definition is as follows:

Let $Z_i$ denote the ith principal component, $i = 1,2,3,...,$ p, Can be set:

$$\begin{cases} Z_1 = c_{11}X_1 + c_{12}X_2 + \ldots + c_{1p}X_p \\ Z_2 = c_{21}X_1 + c_{22}X_2 + \ldots + c_{2p}X_p \\ \cdots \\ Z_p = c_{p1}X_1 + c_{p2}X_2 + \ldots + c_{pp}X_p \end{cases}$$

For each i, all had:

$$c_{i1}^2 + c_{i2}^2 + \cdots + c_{ip}^2 = 1$$

The data are processed based on SPSS platform and principal component analysis algorithm.

- Because these data are the growth rate of bacteria, the original data is no longer standardized to eliminate the dimensional effect.
- The correlation coefficient matrix between variables is established

$$R = (r_{ij})_{m \times m}$$

$$r_{ij} = \frac{\sum_{k=1}^{n} \tilde{x}_{ji} \cdot \tilde{x}_{kj}}{n-1}, (i, j = 1,2, \cdots m)$$

In the formula, $r_{ii} = 1, r_{ij} = r_{ji}, r_{ij}$ Is the correlation coefficient between the i -th index and the j -th index

• Calculate the covariance matrix of 37 * 37, and get the variance percentage of each variable according to the covariance matrix, and then get the cumulative variance value of each component, as shown in Table 1 below

*Table 1: Explanation of total variance*

| Initial eigenvalue | | | | | Extract the load sum of squares | | |
|---|---|---|---|---|---|---|---|
| | component | total | Variance percentage | Cumulative | total | Variance percentage | Cumulative |
| original | 1 | 13.228 | 27.342 | 27.342 | 13.228 | 27.342 | 27.342 |
| | 2 | 9.735 | 20.122 | 47.464 | 9.735 | 20.122 | 47.464 |
| | 3 | 8.089 | 16.720 | 64.184 | 8.089 | 16.720 | 64.184 |
| | 4 | 4.775 | 9.869 | 74.053 | 4.775 | 9.869 | 74.053 |
| | 5 | 2.235 | 4.620 | 78.672 | 2.235 | 4.620 | 78.672 |
| | 6 | 2.132 | 4.406 | 83.078 | 2.132 | 4.406 | 83.078 |
| | 7 | 1.483 | 3.066 | 86.144 | 1.483 | 3.066 | 86.144 |

Based on the statistical analysis software SPSS, the calculation results are shown in the table above, there are seven principal components. The cumulative variance percentage of these seven principal components from small to large is Z1 = 27.342, Z2 = 47.464, Z3 = 64.184, Z4 = 74.053, Z5 = 78.672, Z6 = 83.078, Z7 = 86.144. According to the cumulative variance percentage, the eigenvalue is obtained, and the eigenvalue is under the root sign (cumulative variance percentage), thus the eigenvalue matrix of seven principal components is obtained.

The component score coefficient in the component matrix is to calculate the final value of the final principal component variable, which is also the load matrix of each factor. Based on the component matrix obtained by SPSS platform, the matrix is imported into R-Studio to draw the heat map, as shown in Figure 1

ISSN 2618-1584 Vol. 3, Issue 4: 1-5, DOI: 10.25236/FMSR.2021.030401

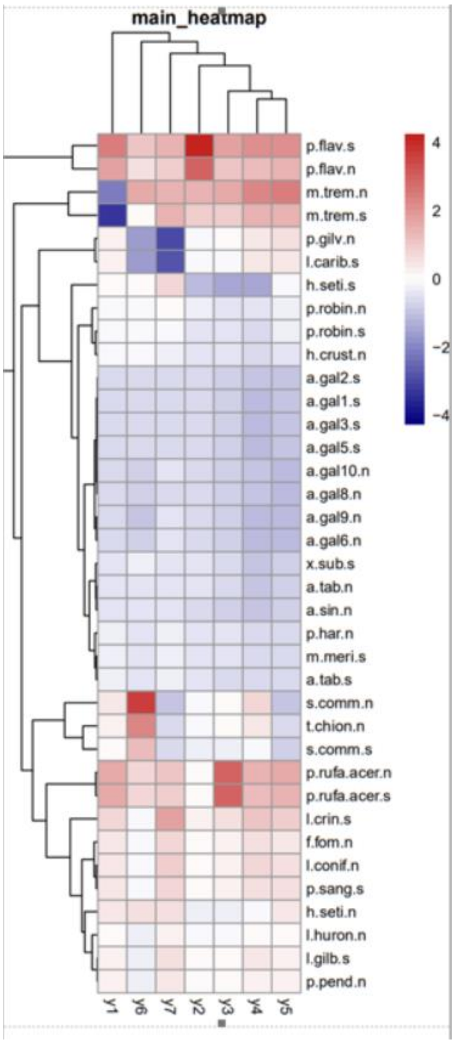


*Figure 1: Heat map of component matrix*

The darker the box color is, the larger the matrix value is, and the greater the contribution in the total contribution rate is.

Finally, the comprehensive score is calculated, and the final comprehensive value of the case = (the value obtained by the principal component variable) * (the corresponding contribution rate). The value obtained by the principal component variable is a 7 * 37 eigenvalue matrix. See the attachment “component-weight.xlsx”. The corresponding contribution rate is the 1 * 7 characteristic matrix mentioned above.

• **Moisture resistance**

The humidity tolerance of fungi is related to the competition ranking of fungi and the difference of water niche width. It is assumed that the greater the difference, the smaller the humidity tolerance; on the contrary, the greater the humidity tolerance.

The niche ranking of each fungus can be obtained. Similarly, competitive ranking can also be ranked, and the difference between water niche ranking and competitive ranking can be made. After standardization, the difference matrix =[1.429,......, -0.131]; Decomposition rate matrix =[0.950, 0.657..., 0.713]; Based on MATLAB platform, the scatter diagram of the two variables of difference value and decomposition rate is drawn as shown in Figure 2:

ISSN 2618-1584 Vol. 3, Issue 4: 1-5, DOI: 10.25236/FMSR.2021.030401

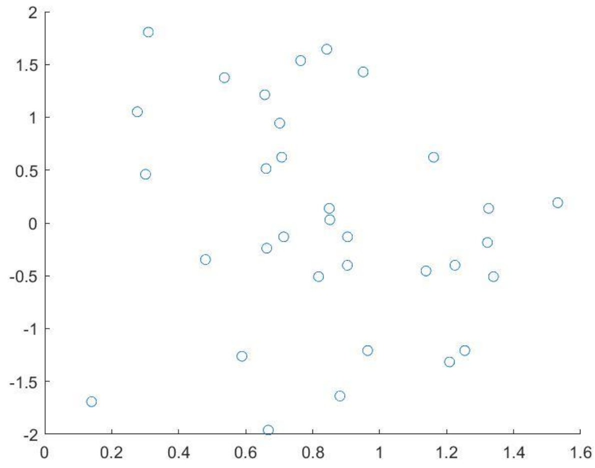

*Figure 2: Scatter plot of difference and decomposition rate*

The final conclusion is that the larger the difference is, the lower the ranking is, the weaker the moisture tolerance is, and the smaller the decomposition rate is. As long as the fungi living environment changes, the decomposition rate will be greatly affected. The smaller the difference, the higher the ranking, the stronger the moisture tolerance, and the lower the impact of environment on the decomposition rate.

We reviewed a large number of literatures and obtained 10 common types of bacteria in the fungal system of saprophytic basidiomycetes. According to the ranking of the 10 types of bacteria and the strength and weakness of the moisture tolerance of bacteria, and the principal component ranking based on the growth rate data of bacteria, the decomposition rate of the two traits of growth rate and moisture tolerance was obtained.

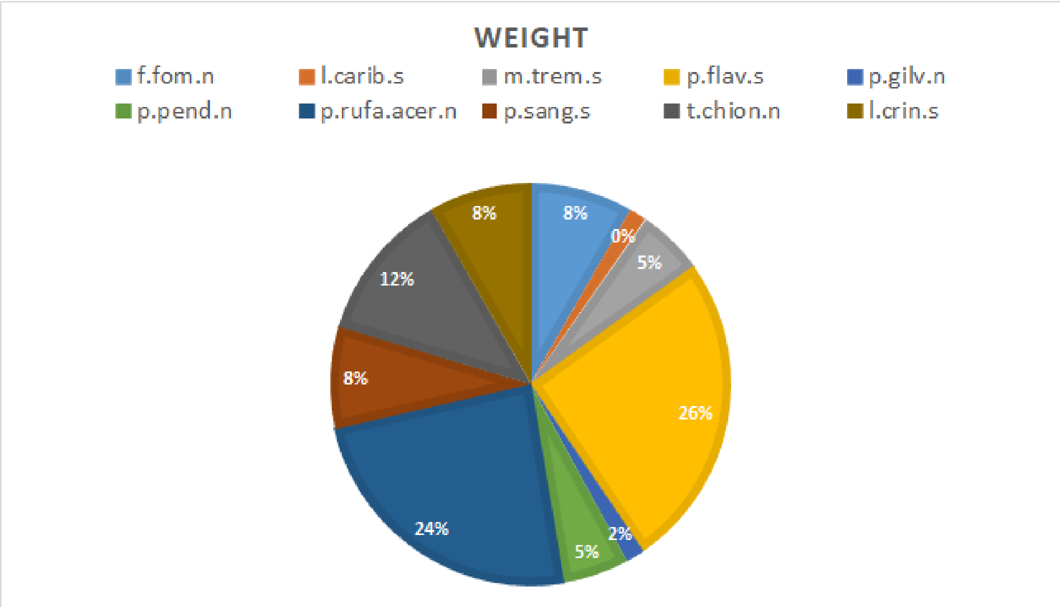


*Figure 3: Principal component weight ratio was analyzed based on common fungi*

There were 10 common types of fungi in the literature fungal interactions reduce carbon use efficiency. According to the weighted sum of the corresponding weight value of these types of bacteria and the decomposition rate of bacteria, the total decomposition rate was obtained under the influence of the two internal characteristics

$$\text{all} - \text{demp} - \text{rate} = -1.64 \times W1 + 6.29 \times W2 + 29.77 \times W3 + 1.8 \times W4 + 6.21 \times W5 + 28.2 \times W6 + 9.55 \times W7 + 14.17 \times W8$$

ISSN 2618-1584 Vol. 3, Issue 4: 1-5, DOI: 10.25236/FMSR.2021.030401

$$\text{Weightvalueof fung} = [-1.64,6.29,29.77,1.8,6.21,28.2,9.55,9.66]$$

$$\text{Bacterialdecompositionrate} = [6.0912,16.008,8.9607,19.8373,1.2768,22.3575,18.5248,9.3353];$$

$$\text{Finalcalculation} = 1298.7$$

## 3. Conclusion

In the short term, the decomposition rate of bacteria is less affected by the external environment, and the external environment changes little, so we build our decomposition rate model based on growth rate and moisture tolerance (differences in water niche and competitive ranking). In this model, the competition result of bacteria is also not considered. In the short term, the same fixed area is full of nutrition, and only the nature and function of bacteria and lignin fiber are considered. We collected literatures to get the growth rate value and decomposition rate value of the pair action of bacteria. Based on the principal component analysis, the large bacteria combination contributing to the overall decomposition rate in the pair action of bacteria combination was used to get the principal component characteristic value. Finally, the comprehensive ranking and weight of all kinds of bacteria were obtained according to the characteristic value. Wet tolerance is plotted in the title to determine how it differs from water niche and competitive ranking. Finally, based on these two properties, PCA algorithm is used again to get the weight value. At this time, the corresponding weight value of each bacterium and the decomposition rate of bacteria are known, and the total decomposition rate of bacteria can be determined finally.

## References

*[1] Venugopal P, Junninen K, Linnakoski R, et al. Climate and wood quality have decayer-specific effects on fungal wood decomposition [J]. Forest Ecology and Management, 2016, 360: 341-351.*

*[2] Silva A S, da Silva A J P, de Carvalho Santos V L. Mathematical model to describe litter decomposition rate in semiarid regions [J]. Revista Brasileira de Ciências Agrárias (Agrária), 2018, 13(4).*

*[3] Lustenhouwer N, Maynard D S, Bradford M A, et al. A trait-based understanding of wood decomposition by fungi [J]. Proceedings of the National Academy of Sciences, 2020, 117(21): 11551-11558.*