Improved estimation model and empirical analysis of relationship between agricultural mechanization level and labor demand

Wang Fulin*, Zhao Shengxue, Fu Xiaoming

(*College of Engineering, Northeast Agriculture University, Harbin 150030, China*)

Abstract: The research on the relationship between mechanization level in planting industry and labor demand was carried out based on the present literatures. The former estimation model of labor demand in planting industry was established without analyzing the effects of planting structure on labor demand in planting industry. The purpose of the research is to develop and perfect the theory of estimating both labor demand and rural surplus labor in planting industry, then to provide some theoretical references for scientific estimation. The model established in this research can be used to calculate not only the amount of current labor demand, but also the demand in the various moment of future according to forecasted mechanization level and cultivated areas. Furthermore, it was explored how to obtain the indexes of cultivated areas, mechanization level and the average cultivated area that each labor can burden when the mechanization level is 0 and 100%. According to statistics principle, the methods of inspection to eliminate abnormal data and data processing were given in order to make the data more credible. Finally, an example was presented for demonstration purposes.

Keywords: agricultural mechanization level, planting industry, labor demand, estimation model, data acquisition, crop production

DOI: 10.3965/j.ijabe.20160902.2188

Citation: Wang F L, Zhao S X, Fu X M. Improved estimation model and empirical analysis of relationship between agricultural mechanization level and labor demand. Int J Agric & Biol Eng, 2016; 9(2): 48－53.

1 Introduction

Labor demand estimation in planting industry is the basis of the estimation of rural surplus labor force[1-3]. The labor demand of a certain region is not only related to the cultivated land area, but also to the level of mechanization[4-7] and the planting structure. Traditional researches only consider the planting area, such as man-days method[8-10], demonstrate method of cultivated land[11,12], base year method[13,14], etc. Wang et al.[15] proposed and established the relationship between the level of mechanization of planting industry and the demand of labor force on the basis of the existing literature researches. However, in this research, the effect of planting structure on the labor demand of the plantation was not considered when establishing the relationship between the level of mechanization and the demand of labor force, which had a certain gap with the actual situation. Because of the difference in the peak period of different crops planting, the land area of a labor planting single crop is less than that of different kinds of crops, but the calculated results would be much more than the amount of labor force needed in actual production[15]. To solve this problem, an improved method was proposed for obtaining and processing the parameters. The research can not only improve the labor demand estimation theory on planting industry, but also lay a theoretical foundation for the scientific estimation of rural surplus labor.

Received date: 2015-07-20 **Accepted date:** 2016-03-18

Biographies: Zhao Shengxue, PhD candidate, Senior Engineer, research interests: agricultural system engineering, Email: zhaoshengxue@163.com; **Fu Xiaoming,** PhD candidate, research interests: agricultural system engineering, Email: dqfxm_1019@163.com.

***Corresponding author: Wang Fulin,** PhD, Professor, research interests: agricultural system engineering. Mailing address: College of Engineering,Northeast Agricultural University, No.59, Mucai Street, Xiangfang District, Harbin 150030, China. Tel: +86-451-55191462, Email: fulinwang1462@126.com.

2 Improved estimation model

According to the definition of mechanization level[16], let $s(t)$ be the area of cultivated land of t moment, $s'(t)$ is the finished area of t moment and $x(t)$ stands for the level of mechanization, then

$$x(t)=\frac{s'(t)}{s(t)}\times 100\% \quad (1)$$

$$s'(t)=s(t)\cdot x(t) \quad (2)$$

Hence, the finished area by labor and livestock $s''(t)$ is

$$s''(t)=s(t)-s'(t)=s(t)[1-x(t)] \quad (3)$$

Let a be the demanded people per hectare with 0 level of mechanization; b is the demanding labor force with 100% level of mechanization; $l(t)$ means the demanded people of planting industry at time t. Mechanization level 0 refers to all operations are done by human and animal; Mechanization level 100% refers to all assignments are complete by machinery. The following mathematical model can be established:

$$L(t)=as(t)[1-x(t)]+bs(t)x(t) \quad (4)$$

where, the first items denoted the amount of labor of people and living stocks, and the second item denoted the amount of labor when using machines.

In the actual calculation, we can use the index of cultivated land area per capita burden in order to facilitate data acquisition, so the relationship between the amount of labor force needed for crops per hectare and cultivated land area per capita burden can be described as:

$$a=\frac{1}{A} \quad (5)$$

$$b=\frac{1}{B} \quad (6)$$

where, A is the planting area burden by per labor with 0 level of mechanization, hm^2; B is the planting area with 100% level of mechanization, hm^2.

Put Equations (5) and (6) into Equation (4), then,

$$L(t)=\frac{s(t)}{A}+\left[\frac{s(t)}{B}-\frac{s(t)}{A}\right]x(t) \quad (7)$$

The Equation (7) is the improved model for the estimation of labor demand in plantation presented in this research. In this model, A and B are values of aggregative indicators which take in the account of many factors such as regional characteristics, planting structure, natural conditions, cropping patterns, meteorological factors, etc. The key is the confirmation of values of A and B when putting it into use of this model.

If only taking the demand of labor force into account, the static model is:

$$L=\frac{s}{A}+\left[\frac{s}{B}-\frac{s}{A}\right]x \quad (8)$$

where, L is the amount of demand for the labor force; s is planting area, hm^2; x is the level of mechanization, %.

Using the model described in Equations (7) and (8), we can calculate the amount of labor demand in a certain area and forecast the amount of labor demand in anywhere and anytime in the future.

3 Data acquisition and processing

3.1 Data acquisition

Cultivated area, the degree of mechanization, the cultivated land area with 0 or 100% level of mechanization are included in the data acquisition of labor force demand in the estimation model. Those are $s(t)$, $x(t)$, A, B in Equation (7). After analysis, $s(t)$ and $x(t)$ belong to the same class of data. A and B belong to the same class of data.

3.1.1 Data acquisition of planting area and the level of mechanization

When calculating labor demand in one area at certain moment, we need to estimate according to the historical data of planting area and the level of mechanization which can be acquired by statistical yearbook. In general, the planting area and level of mechanization can be estimated by regression model and Delphi method[17].

3.1.2 Acquisition of affordable planting area per labor

The data for the affordable planting area per labor when the level of mechanization between 0 and 100% are unavailable from statistical yearbook, but these can be obtained from the investigations of actual actives of plantings.

There are two methods to acquire A and B. One is to investigate from the farmers, another is to survey of farmers in the towns and villages of the different degree of mechanization. Using this method, we can get the average cultivated land area per capita, then the value of A and B can be calculated using the regression model.

No matter which method is chosen, the survey sample must be sufficient, generally should not less than 50 samples, and it must be statistical test. If the statistical test does not pass, it should expand the number of samples, until the statistical test satisfied.

During the survey, first is to prepare the survey table; the second is to make it clear that the cultivated land per capita burden is decided by the busiest planting time, that is, the maximum number of labor input; the third is that hired labor should be calculated in the amount of labor demand, if there is a necessary of seasonal workers.

3.2 Data processing

3.2.1 Distribution test

In order to improve the reliability of data, it is necessary to test if there is abnormal data, and the abnormal data should be removed. The choice of data eliminating method is determined by the distribution of data. Therefore, it is necessary to test the distribution of the data.

The common method of distribution test is χ^2 test, the basic idea is: All the results of the random test are divided into A_1, A_2, ..., A_K ($A_1 \cup A_2 \cup ... \cup A_K = \Omega$, $A_i \cap A_j = \varnothing$, $i \neq j$; i, j=1, 2, ..., k), So under the assumption H_0, it can calculate $p_i=P(A_i)$, i=1, 2, ..., k.

Obviously, in n trials, frequency of events f_i/n differences with the p_i, Generally, if H_0 stand, the difference was not significant and if H_0 does not exist, the difference is significant. Based on the idea, Pearson is used as statistic

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - np_i)^2}{np_i} \tag{9}$$

As a test hypotheses, H_0 tallies with the actual size.

The actual step of χ^2 test method is that let general ξ theoretical distribution is $F(x)$, $F_0(x)$ is given in advance of a distribution function. In order to test:

$$H_0 : F(x) = F_0(X) \quad H_1 : F(x) \neq F_0(x) \tag{10}$$

1) The sample value range is divided into k mutually disjoint intervals:

$(a_0, a_1]$, $(a_1, a_2]$, $\cdots$, $(a_{i-1}, a_i]$, $\cdots$, $(a_{k-1}, k_k]$, and $-\infty < a_0 < a_1 < \cdots < a_{i-1} < a_i < a < \cdots < a_k < a_{k-1} < +\infty$.

2) Calculating the sample values falling into each range $(a_{i-1}, a_i]$ (i=1, 2, ..., k).

3) Under the theory of distribution ξ falling into $(a_{i-1}, a_i]$, when H_0 is true, the probability falling into $(a_{i-1}, a_i]$ is:

$$p_i = P\{a_{i-1} < \xi \leq a_i\} = F_0(a_{i-1}), (i = 1, 2, \cdots, k) \tag{11}$$

Hence we get ξ falling into $(a_{i-1}, a_i]$, this theoretic frequency np_i, np_i should not be less 5.

4) Statistical magnitude:

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - np_i)^2}{np_i}$$

According to Pearson theory, $\chi^2 \sim \chi^2(k - r - 1)$, r is the amount being estimated in $F_0(x)$.

5) Given level of significance α, check the χ^2 distribution table, we can have threshold $\chi_\alpha{}^2(k - r - 1)$.

6) Calculated χ^2 from sample observation value of the statistic.

7) Induce: when $\chi^2 > \chi_\alpha{}^2(k - r - 1)$, reject H_0, that is, $F(x)$ unmatches $F_0(x)$. when $\chi^2 < \chi_\alpha{}^2(k - r - 1)$, accept H_0, $F(x)$ matches $F_0(x)$.

χ^2 test method is derived when n is infinite, so must pay attention to use N that is large enough, and np_i is not so small. According to experience, the required sample size n is not less than 50, and each np_i is not less than 5, and np_i is best above 10, otherwise should be properly merged interval (or A_i), make the np_i to satisfy this requirement.

3.2.2 Elimination of abnormal data

Suppose that the samples of the affordable planting area per labor be confirming to normal distribution $N(\mu, \sigma^2)$, μ is the mathematical expectation of samples of affordable planting area per labor, σ^2 is the variance of affordable planting area per labor, according to the principle of mathematical statistics[18], if probability density of random variable ξ is:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(x-\mu)^2}{2\sigma^2}}, \quad (-\infty < x < +\infty) \tag{12}$$

where, $-\infty < x < +\infty$, $\sigma > 0$ is an constant, then ξ obeys the normal distribution with parameters of μ and σ, i.e. $\xi \sim N(\mu, \sigma^2)$. The function of Normal Distribution random variable ξ is:

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dt \tag{13}$$

Specially, when μ=0, σ=1, the normal distribution is standard normal distribution $N(0,1)$, its probability density is:

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2}},\ (-\infty<x<+\infty) \tag{14}$$

The distribution function is:

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt \tag{15}$$

For a random variable of normal distribution, such as $\xi\sim N(\mu, \sigma^2)$, the can get the probability by using $u=\frac{x-\mu}{\sigma}$ and $\Phi(u)$, since:

$$\begin{aligned}F(x)&=\int_{-\infty}^{x}f(t)dt=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx\\&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{x-\mu}{\sigma}}e^{-\frac{u^2}{2}}du=\Phi\left[\frac{x-\mu}{\sigma}\right]\end{aligned} \tag{16}$$

According to the principle shown above, the probability variance σ those three times of absolute difference of an inspect value ξ and mathematical expectation (average) is:

$$P\{|\xi-\mu|<3\sigma\}=\frac{1}{\sqrt{2\pi}\sigma}\int_{\mu-3\sigma}^{\mu+3\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx=0.9974 \tag{17}$$

While it varies of standards for different countries to control quality, it is bounded by $\mu\pm3\sigma$ generally, that is, considering it's normal when the inspect values fall into [$\mu\pm3\sigma$], and abnormal when miss and should be eliminated, hence we call this elimination principle the 3σ Principle.

3.2.3 The method of data arrangement

Suppose the inspect sample of affordable planting area per labor A is N_1 when the level of mechanization is 0, and the affordable planting area per labor b is N_2, after inspection if the number of A's abnormal data is $m_1(m_1<N_1)$ and B's $m_2(m_2<N_2)$, let the number of A's inspection value be N_1' and B's be N_2', we have:

$$N_1'=N_1-m_1 \tag{18}$$

$$N_2'=N_2-m_2 \tag{19}$$

Let the i^{th} sample be A_i of affordable planting area per labor after eliminating abnormal samples and the level of mechanization is 0, and the j^{th} sample is B_j when level of mechanization is 100%, we have:

$$A=\frac{1}{N_1'}\sum_{i=1}^{N_1'}A_i \tag{20}$$

$$B=\frac{1}{N_2'}\sum_{j=1}^{N_2'}B_j \tag{21}$$

According to method of data arrangement above, A and B are believable.

4 Demonstration of calculation

There are 640 000 hm^2 farm land in northern of Heilongjiang Land Reclamation Area, the calculated level of mechanization is 82% after eliminating abnormal data, as shown in Table 1.

Table 1 Cultivated area, level of mechanization, the cultivated land of crops per labor when the level of mechanization is 0 and 100% in northern of Heilongjiang Land Reclamation Area

Cultivated area/hm^2	Level of mechanization/%	Cultivated land area of crops per labor with 0 level of mechanization/hm^2	Cultivated land area of crops per labor with 100% level of mechanization/hm^2
640 000	82	1.5	6.6

The demand of labor force (L) in this area can be obtained after putting cultivated land area, level of mechanization into Equation (8):

$$\begin{aligned}L&=\frac{s}{A}+\left[\frac{s}{B}-\frac{s}{A}\right]x\\&=\frac{1620000}{1.5}+(\frac{1620000}{6.6}-\frac{1620000}{1.5})\times0.82\\&=412517\end{aligned}$$

The level of mechanization and change of farm land area are shown in Table 2 according to Heilongjiang statistical yearbook. We have the Regression model after the regressing of level of mechanization of 2013-2014:

$$\hat{x}(t)=\frac{1}{1+0.757e^{-0.112t}} \tag{22}$$

in which the correlation coefficient r=0.996>r(0.05)and inspection value F=1010.618>F(0.05). t is time variable, t=1, 2, …, 10 is denoting to year 2004, 2005, … , 2013. The regression model is as follows:

$$\hat{s}(t)=1442540.90-\frac{175526.30}{t} \tag{23}$$

in which correlation coefficient r=0.989>r(0.05), and the inspect value F=369.249>F(0.05).

According to Equations (22) and (23), the results are obtained after a forecast of the level of mechanization and farm land area, and then get the demand of labor force. The forecasting value can be drawn as Table 3.

Table 2 Mechanization level and cultivated land in northern of Heilongjiang Land Reclamation Area from 2004-2013

Year	Level of mechanization/%	Cultivated land area/hm^2
2004	60.0	1 610 000
2005	62.6	1 544 000
2006	65.1	1 505 330
2007	67.5	1 496 000
2008	69.9	1 474 400
2009	72.1	1 473 370
2010	74.3	1 461 140
2011	76.3	1 460 000
2012	78.2	1 458 080
2013	80.1	1 457 200

Table 3 Forecasting value of mechanization level, cultivated land, plant industry labor demand quantity from 2014 to 2024

Year	Level of mechanization/%	Cultivated land area/hm^2	Labor demand /person
2014	81.9	1 458 497.8	356 979
2015	83.5	1 457 168.1	344 643
2016	85.0	1 456 042.9	333 200
2017	86.4	1 455 078.5	322 635
2018	87.6	1 454 242.7	313 011
2019	88.8	1 453 511.3	304 100
2020	89.9	1 452 866.0	295 981
2021	90.8	1 452 292.4	288 567
2022	91.7	1 451 779.1	281 807
2023	92.5	1 451 317.2	275 656
2024	93.3	1 450 899.3	270 067

5 Conclusions

1) This study points out the flaws of former researches on relationship between level of mechanization and demand of labor force which did not consider the effect of planting structure to demand of labor force.

2) An improved model of relationship between level of mechanization and labor demand under planting structure was proposed. It can not only calculate the demand for certain farm land area, but also can carry out the labor demand of every moment by forecasting future level of mechanization and farm land area.

3) The new method was proposed for data acquisition of farm land area, level of mechanization and affordable planting area per labor. At the meantime, we put forward the relevant problems to be considered in obtaining farm land area per labor.

4) According to the principle of statistical, we studied and proposed methods of elimination, inspection, and arrangement for farm land area per labor when level of mechanization is 0 and 100%.

5) Taking the northern of Heilongjiang Land Reclamation Area as an example, the relationship between level of mechanization and demand of labor force was calculated that both suitable for the present and future.

Acknowledgement

This research was supported by the National Natural Science Foundation of China (31071331) and National Social Science Foundation of China (13BJY098).

[References]

[1] Yi B. The Prediction on the Scale of Rural Surplus Labor Transfer in Urban Planning: A Case Study of Kunming. Urban Insight, 2015.

[2] Hone N. The research on transfer employment of rural surplus labor force in the process of the new-type urbanization. International Journal of Paediatric Dentistry, 2010; 20(3): 201–206.

[3] Ren J, Zhang G. Analysis of the Employment of Rural Surplus Labor Force in Cities. Agricultural Science & Technology & Equipment, 2014; 21(3): 24–29.

[4] Zhang D P, Shi G D, Chen J G. Agricultural mechanization’s effect on agricultural labor surplus. Transactions of the Chinese Society of Agricultural Machinery, 1996; 27(4): 1–6. (in Chinese with English abstract)

[5] Lambeth R M. Scabbing the Palouse: agricultural labor replacement and union busting in southeast Washington, 1917–1919. Labor History, 2015; 56(4): 1–20.

[6] Lou J, Zhu H J, Han L J. Relationship between agricultural labor transfer and mechanization. System Sciences and Comprehensive Studies in Agriculture, 2006; 22(1): 66–69. (in Chinese with English abstract)

[7] Dong Q, Murakami T, Nakashima Y. The Recalculation of the agricultural labor forces in China. Social Science Research Network, 2015. Available at SSRN: http://ssrn.com/abstract=2630513 or http://dx.doi.org/10.2139/ssrn.2630513

[8] Juhn C, Potter S. Changes in Labor Force Participation in the United States. Journal of Economic Perspectives, 2006; 20(3): 27–46.

[9] Chen H. Explore of the new path of the rural surplus labor

transfering. Transfusion Medicine Reviews, 2015; 34(4): S285–S286.
[10] Suo R X, Wang F L, Sun Xu. Study on labor-days method of estimation rural surplus labor force analysis and improvement. Mathematics in Practice and Theory, 2011; 41(3): 108–112. (in Chinese with English abstract)
[11] Chaozhou L U, Luo Y. The estimation methods for agricultural surplus labor based on stochastic Fourier production function. Asian Agricultural Research, 2014; 6: 41–48.
[12] Miao W W, Zhang L X, Guo L N. Surplus labor estimation of agricultural sector based on view of Marx's Necessary Labor and Surplus Labor. Journal of Anhui Agricultural Sciences, 2014. (in Chinese with English abstract)
[13] Wu Q J, Yang Y L, Yang D. A study on the mathematical model of rural surplus labor. Commercial Modernization, 2006; 489: 25–26.
[14] Pham Q D, Vo H H. Measurement of Surplus Labor in Viet Nam Agriculture. Journal of Mathematics and System Science, 2014; 2: 105–110.
[15] Wang F L, Suo R X, Zhang L, Cang L. Study on the relationships between plant industry mechanization extent and the demand for labor. Transactions of the CSAE, 2010; 26(9): 181–184. (in Chinese with English abstract)
[16] Yu Y T. Agricultural Mechanization Engineering. China Outlook Press, Beijing, 1987. (in Chinese)
[17] Ameyaw E E, Hu Y, Shan M, et al. Application of Delphi method in construction engineering and management research: a quantitative perspective. Journal of Civil Engineering & Management, 2016; 42(2): 1–10.
[18] Rice J A. Mathematical Statistics and Data Analysis. USA: Duxbury Resource Center, 2006.