Establishment and verification of labor demand estimation model in planting industry

Xu Bing, Wang Fulin*, Wang Jiquan, Zhao Shengxue

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

Abstract: In view of the poor precision of the theoretical model of labor demand estimation, it is difficult to estimate and predict the actual production problems accurately.   Based on the actual production conditions and the relationship between the degree of mechanization of planting and the demand of labor force, this study established an estimation model for the labor demand of planting industry considering the factors of planting structure and mechanization degree.   In order to ensure high reliability of data, the method of checking out abnormal data was adopted to obtain the cultivated land area index when the mechanization degree is from 0 to 100%.   Taking Suihua region (Heilongjiang Province, China) as an example, the theory of the research was analyzed and applied.   This study accessed to the data of cultivated land area per labor can afford when the mechanization level in Suihua area were 0 and 100% respectively through the investigation, and the average cultivated land area data of each labor force in two cases were sorted out and the abnormal data were eliminated at the same time.   Finally, using the derived model, the data obtained and the mechanization level and cultivated land area of Suihua in the future, the labor demand amount in Suihua area from 2015 to 2019 were predicted.   The model established in this study can be used to calculate the quantity of both current labor demand in planting industry and the labor demand in the various moments in the future through forecasting the future mechanization level and cultivated area which are the two main factors influencing the quantity of labor demand in planting structure.

Keywords: planting industry, labor, estimation model, case verification, forecast evaluation

DOI: 10.25165/j.ijabe.20171006.3094

1   Introduction

The problem of estimating labor demand in planting industry is the basis of the estimation of the rural surplus labor[1-4].   The factors influencing an area labor demand in planting industry are the amount of cultivated land area, the degree of the planting mechanization[5-7], the diversity of planting structure, etc.   To solve this issue, traditional researches are only considering the cultivated land area.   The Labor-Days Method calculates the amount of labor demand through the per hectare labor-days demand, amount of labor-days can be provided per labor per year and the cultivated land area[8-12].   Per-capita Arable Area Method is a method to estimate the demand of agricultural labor based on the area of arable land.   The key point of this method is to determine the per capita arable area[13,14].   Base Year Method selects one year as base year according experience.   In the selected year, the labor supply balances with labor demand.   The per-capita arable area can be determined through the arable land and amount of agricultural labor.

Then all the other years' labor demand can be calculated compared with the base year[15,16].

Based on the existing literature research, Wang et al.[17,18] first proposed and established the relation model between the degree of mechanization and the labor demand in planting industry. However, the relation model did not consider the influences of the planting structure on the labor demand in planting industry. There was a certain gap between the model and the actual production. To sum up, the calculated results of the model were bigger than the actual labor demand in planting industry.

In order to solve this problem, an improved calculation model was presented and the acquisition and processing method of model parameters were also addressed. The research was expected to provide a reference for scientific and accurate estimation of the surplus rural labor force in planting industry.

2 Establishment of the labor demand estimation model in planting industry

Based on the definition of the degree of mechanization[19], assuming that $s(t)$ is the cultivated area at time t, $s'(t)$ is the area cultivated by machines at time t and $x(t)$ is the degree of mechanization at time t.   Then,

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

$$s'(t) = s(t) \cdot x(t) \tag{2}$$

So the area cultivated by human and animal power at time t is:

$$s''(t) = s(t) - s'(t) = s(t)[1 - x(t)] \tag{3}$$

Using a and b to express the human number of labor demand per hectare cultivated with the degree of mechanization of 0 or 100%, respectively, $L(x(t))$ to express the human number of labor demand in planting industry at time t , thus, the mathematical model can be established as follows:

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

where, the first item expresses the labor demand of the cultivation utilizing human power and animal power. And the second item expresses the labor demand of the mechanized cultivation.

Defining $l(x(t))$ as the cultivated land that the average labor covered at time t when the degree of mechanization is $x(t)$, the equation can be expressed as:

$$l(x(t)) = \frac{s(t)}{L(x(t))} \tag{5}$$

Substituting Equation (4) into Equation (5), the $l(x(t))$ can be expressed as Equation (6):

$$l(x(t)) = \frac{s(t)}{as(t)[1 - x(t)] + bs(t)x(t)}$$
$$= \frac{1}{a[1 - x(t)] + bx(t)} \tag{6}$$

In the practical application of the model above, the index of per capita affordable arable area can be used for convenience of data acquisition.   For $l(x(t))$ is defined as the cultivated land that the average labor covered at time t when the degree of mechanization is $x(t)$.   When the degree of mechanization is 0 at any time t, $x(t)$ always equals 0 and $l(x(t))$ is $l(0)$.   Similarly, when the degree of mechanization is 100% at any time t, $x(t)$ always equals   100% and $l(x(t))$ is $l(100\%)$.

Since the cultivated land that the average labor covered at time t with 0 degree of mechanization is:

$$l(0) = \frac{1}{a} \tag{7}$$

And the cultivated land that the average labor covered at time t with 100% degree of mechanization is:

$$l(100\%) = \frac{1}{b} \tag{8}$$

Substituting Equations (7) and (8) into Equation (6), Equation (9) can be obtained as follows:

$$l(x(t)) = \frac{1}{\frac{1}{l(0)} \cdot (1 - x(t)) + \frac{1}{l(100\%)} \cdot x(t)}$$
$$= \frac{l(100\%) \cdot l(0)}{l(0) \cdot x(t) + l(100\%) \cdot (1 - x(t))} \tag{9}$$

Equation (9) is the relation model between the degree of mechanization and the average labor force can cover the area of cultivated land in planting industry.   There are two parameters $l(100\%)$ and $l(0)$ in this model, from Equation (5), $L(x(t))$ can be expressed as:

$$L(x(t)) = \frac{s(t)}{l(x(t))}$$
$$= \frac{l(0) \cdot x(t) + l(100\%) \cdot (1 - x(t))}{l(100\%) \cdot l(0)} \cdot s(t) \tag{10}$$

The Equation (10) is the improved estimation model of labor demand in planting industry in this article. In this model $l(100\%)$ and $l(0)$ are comprehensive indexes considering regional characteristics, meteorological factor, natural conditions, cultivation model, planting construction and some other factors. The key point to apply this model is to determine the values of $l(100\%)$ and $l(0)$.

If only to calculate the current labor demand in planting industry, the static model can be obtained as:

$$L(x) = \frac{l(0) \cdot x + l(100\%) \cdot (1-x)}{l(100\%) \cdot l(0)} \cdot s \qquad (11)$$

where, $L(x)$ is the number of labor demand in planting industry; s is the cultivated area; x is the degree of mechanization.

Using Equations (10) and (11) we can estimate the number of current labor demand in planting industry on a particular area. These two equations can be also used to estimate the labor demand at any time in the future through forecasting the future mechanization degree in particular area at particular time.

3   Typical case verification

3.1   Research status and methods

Twelve villages were selected randomly in Suihua, Heilongjiang, China as the survey sites, questionnaires were collected from 120 farmers. Defining N_1 as the survey sample size of per capita affordable arable area $l(0)$ with 0 degree of mechanization, N_2 as the survey sample size of per capita affordable arable area $l(100\%)$ with 100% degree of mechanization, m_1 $(m_1 < N_1)$ as the number of abnormal data in the sample of $l(0)$, m_2 $(m_2 < N_2)$ as the number of abnormal data in the sample of $l(100\%)$, and N'_1 and N'_2 as the survey sample size of the cultivated land that the average labor covered without the abnormal data for $l(0)$ and $l(100\%)$, respectively, the equations can be given as:

$$N'_1 = N_1 - m_1 \qquad (12)$$

$$N'_2 = N_2 - m_2 \qquad (13)$$

After excluding the abnormal data in the sample, if the cultivated land of sample f when the degree of mechanization is 0 is $l(0)_f$, $f=1,2,\ldots, N'_1$. The cultivated land of sample g when the degree of mechanization is

100% is $l(100\%)_g$, $g=1,2,\ldots, N'_2$. The equations can be given as:

$$l(0) = \frac{1}{N'_1} \sum_{f=1}^{N'_1} l(0)_f \qquad (14)$$

$$l(100\%) = \frac{1}{N'_2} \sum_{g=1}^{N'_2} l(100\%)_g \qquad (15)$$

To process data by the method above, the cultivated land that the average labor covered $l(0)$ and $l(100\%)$ can be identified as credible.

3.2   Acquisition of production data

In the estimation model of labor demand in planting industry, there are two model parameters $l(0)$ and $l(100\%)$. These two parameters cannot be obtained directly from statistical information, which can be obtained from the actual survey data. First a survey of famers in a particular area was conducted to get the data of the cultivated land that the average labor covered under the current degree of mechanization. Then the regression model was established with the degree of mechanization and per capita affordable arable area to calculate the values of $l(0)$ and $l(100\%)$.

The acquisition of the sample should be larger than 50. Statistical tests should be done with survey sample. If the test does not pass, samples need to be further increased until the test passed[20].

The attention should be paid to following three aspects when doing surveys. Firstly, survey forms should be well-designed in advance. Secondly, it should be clear stated that amount of average labor covered cultivated land is determined by the peak of cultivation. Thirdly, the seasonal labor demand should be included in the labor demand in planting industry.

3.3   Data distribution test

In order to make the data from survey more reliable, it is necessary to do distribution test of survey data. If there are some abnormal data, they should be removed from survey sample. The methods to eliminate the abnormal data are determined by regularities of distribution of the survey sample.

The common method to do distribution test of survey data is chi-square test[21]. If all the results of the survey are expressed by Ω. Divide Ω into k no related events.

They are shown as A_1, A_2, $\cdots$, $A_K (A_1 \cup A_2 \cup \cdots \cup A_K = \Omega)$. Among them, $A_i \cap A_j = \varnothing$, $i \neq j$ and i, $j=1,2,\ldots,k$. Hypothesis is formulated as H_0. It can calculate $p_i = P(Ai)$, $i=1,2,\ldots,k$.. Obviously, when the survey sample is n, there are some differences between the frequency (f_i/n) of the survey result Ai and p_i. Generally, if the hypothesis H_0 is established, the difference between f_i/n and p_i is not significant. Otherwise, the difference between f_i/n and p_i is significant. Therefore, the statistics shown as Equation (16) are used as a measure of the degree of agreement between the hypothesis H_0 and the physical truth[22-24].

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} \qquad (16)$$

According to Pearson theorem, it can be seen that χ^2 obeying $\chi^2(k-r-1)$ distribution. r is the number of estimated parameters in $F_0(x)$. The critical value $\chi_\alpha^2(k-r-1)$ can be obtained through looking up the distribution table of χ^2 when the significance level α is given. The value of χ^2 can be calculated by the survey sample. When $\chi^2 > \chi_\alpha^2(k-r-1)$, the hypothesis H_0 should be rejected. When $\chi^2 < \chi_\alpha^2(k-r-1)$, the hypothesis H_0 should be accepted. n tends to infinity in the chi-square test. Accordingly, the sample size np_i should be large enough when using chi-square test. np_i cannot be too small as well. Generally, the sample size should be larger than 50 and all the value of np_i should be no less than 5. Otherwise, intervals should be merged.

3.4 Data processing method

If the cultivated land that the average labor covered is in the normal distribution $N(\mu,\sigma^2)$ through testing, μ is the excepted value of the cultivated land that the average labor covered and σ^2 is the variance of the cultivated land

that the average labor covered, according to probability and mathematical statistics[25,26].

Generally, standard normal distribution is the normal distribution when $\mu=0$, $\sigma=1$. It is shown as $N(0,1)$. Its probability density and distribution function is:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2}} \quad (-\infty < x < +\infty) \qquad (17)$$

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt \qquad (18)$$

By transforming with $u = \dfrac{x-\mu}{\sigma}$, looking up the table by using the value of $\Phi(u)$, then,

$$F(x) = \int_{-\infty}^{x} f(t)dt = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{u^2}{2}} du = \Phi\left[\frac{x-\mu}{\sigma}\right] \qquad (19)$$

According to the principles above, the probability of absolute value of the difference between ξ and mathematical expectation μ (mean value) less than 3 is:

$$P\{|\xi - \mu| < 3\sigma\} = \frac{1}{\sqrt{2\pi}\sigma} \int_{\mu-3\sigma}^{\mu+3\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 0.9974 \quad (20)$$

The majority regards $\mu \pm 3\sigma$ as the boundary, if the observed value fall into the interval $[\mu+3\sigma]$, the data will be normal. Otherwise, the data will be abnormal and should be eliminated[27].

3.5 Results processing and analysis

The sample data of the cultivated land that the average labor covered with 0 and 100% degree of mechanization were obtained. Sample data are shown in the following Tables 1 and 2 after sorting. By using Excel software, we can make the frequency histogram of the survey data of the cultivated land area when the mechanization degree is 0 and 100%, as shown in Figures 1 and 2.

Table 1   Sample data of per capita affordable arable area with 0 degree of mechanization (hm²/person)

0.8667	1.1333	1.2000	1.2667	1.3333	1.3333	1.4000	1.4000	1.4667	1.4667	1.5333	1.6667
0.9333	1.1333	1.2000	1.2667	1.3333	1.3333	1.4000	1.4000	1.4667	1.4667	1.5333	1.6667
1.0000	1.1333	1.2000	1.2667	1.3333	1.3333	1.4000	1.4000	1.4667	1.5333	1.6000	1.6667
1.0000	1.1333	1.2667	1.2667	1.3333	1.3333	1.4000	1.4667	1.4667	1.5333	1.6000	1.6667
1.0667	1.1333	1.2667	1.2667	1.3333	1.3333	1.4000	1.4667	1.4667	1.5333	1.6000	1.6667
1.0667	1.2000	1.2667	1.2667	1.3333	1.4000	1.4000	1.4667	1.4667	1.5333	1.6000	1.7333
1.0667	1.2000	1.2667	1.2667	1.3333	1.4000	1.4000	1.4667	1.4667	1.5333	1.6000	1.8000
1.0667	1.2000	1.2667	1.2667	1.3333	1.4000	1.4000	1.4667	1.4667	1.5333	1.6000	1.8667
1.0667	1.2000	1.2667	1.3333	1.3333	1.4000	1.4000	1.4667	1.4667	1.5333	1.6000	2.0000
1.1333	1.2000	1.2667	1.3333	1.3333	1.4000	1.4000	1.4667	1.4667	1.5333	1.6667	2.0000

Table 2   Sample data of per capita affordable arable area with 100% degree of mechanization (hm²/person)

|											
6.3333	7.3333	7.6667	7.6667	7.6667	8.0000	8.0000	8.0000	8.3333	8.3333	8.3333	8.6667
6.6667	7.3333	7.6667	7.6667	7.6667	8.0000	8.0000	8.0000	8.3333	8.3333	8.3333	8.6667
6.6667	7.3333	7.6667	7.6667	7.6667	8.0000	8.0000	8.0000	8.3333	8.3333	8.3333	8.6667
7.0000	7.3333	7.6667	7.6667	7.6667	8.0000	8.0000	8.0000	8.3333	8.3333	8.6667	9.0000
7.0000	7.3333	7.6667	7.6667	8.0000	8.0000	8.0000	8.0000	8.3333	8.3333	8.6667	9.0000
7.0000	7.3333	7.6667	7.6667	8.0000	8.0000	8.0000	8.0000	8.3333	8.3333	8.6667	9.0000
7.0000	7.3333	7.6667	7.6667	8.0000	8.0000	8.0000	8.0000	8.3333	8.3333	8.6667	9.0000
7.3333	7.3333	7.6667	7.6667	8.0000	8.0000	8.0000	8.3333	8.3333	8.3333	8.6667	9.3333
7.3333	7.6667	7.6667	7.6667	8.0000	8.0000	8.0000	8.3333	8.3333	8.3333	8.6667	9.6667
7.3333	7.6667	7.6667	7.6667	8.0000	8.0000	8.0000	8.3333	8.3333	8.3333	8.6667	9.6667
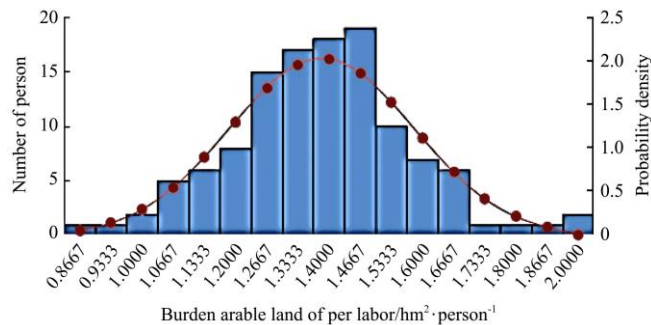
Figure 1   Frequency distribution histogram of burden arable land of per labor when mechanization degree is 0
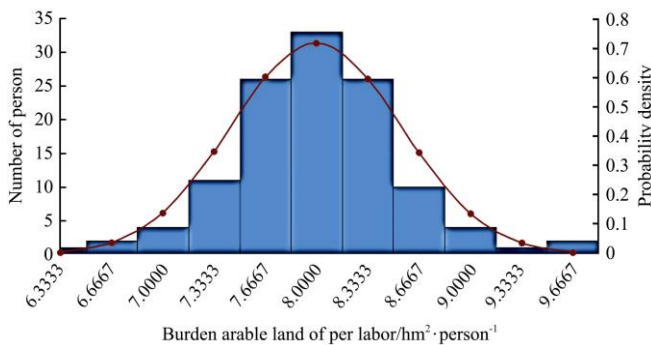

Figure 2   Frequency distribution histogram of burden arable land of per labor when mechanization degree is 100%

The mean value, variance and standard deviation of the sample can be calculated separately as follows:

$$\hat{\mu}(0) = \bar{X}(0) = 1.3850$$

$$\hat{\sigma}^2(0) = \frac{1}{n}\sum_{i=1}^{n}[X_i(0) - \bar{X}(0)]^2 \approx 0.0385$$

$$\hat{\sigma}(0) \approx 0.1962$$

According to the data of per capita affordable arable area when the degree of mechanization is 100% in the Table 2, the mean value, variance and standard deviation of the sample can be calculated separately as follows:

$$\hat{\mu}(100\%) = \bar{X}(100\%) = 7.9972$$

$$\hat{\sigma}^2(100\%) = \frac{1}{n}\sum_{i=1}^{n}[X_i(100\%) - \bar{X}(100\%)]^2 \approx 0.3065$$

$$\hat{\sigma}(100\%) \approx 0.5536$$

The intervals of the normal distribution function is $(-\infty, +\infty)$. At this time if the survey data in the Tables 1 and 2 all obey normal distribution, these two tables can be both divided into several disjoint intervals. In this case study, Table 1 is divided into 9 disjoint intervals. The first interval and the last one are $(-\infty, 1.01]$ and $(1.71, +\infty)$, respectively. The rest intervals are divided by group distance of 0.1. Table 2 is also divided into 9 disjoint intervals. The first interval and the last one are $(-\infty, 6.92]$ and $(9.02, +\infty)$, respectively. The rest intervals are divided by group distance of 0.3.

If the hypothesis H_0 is established, the probability of survey data ξ in Table 1 representing the cultivated land that the average labor covered when the degree of mechanization is 0 falling into every interval is:

$$p_i(0) = P(0)\{a_{i-1}(0) < \xi(0) \le a_i(0)\}$$

$$= \Phi(\frac{a_i(0) - 1.3850}{0.1962}) - \Phi(\frac{a_{i-1}(0) - 1.3850}{0.1962}) \quad (21)$$

The probability of survey data ξ in Table 2 representing the cultivated land that the average labor covered when the degree of mechanization is 100% falling into every interval is:

$$p_i(100\%) = P(100\%)\{a_{i-1}(100\%) < \xi(100\%) \le a_i(100\%)\}$$

$$= \Phi(\frac{a_i(100\%) - 7.9972}{0.5536}) - \Phi(\frac{a_{i-1}(100\%) - 7.9972}{0.5536}) \quad (22)$$

When $i=1,2,\ldots,k$, after proper interval merging, $p_i(0)$ and $p_i(100\%)$ can be calculated. The results are shown in Tables 3 and 4.

When the degree of mechanization is 0, the value of χ^2 is:

$$\chi^2 = \sum_{i=1}^{8}\frac{[n_i - np_i(0)]^2}{np_i(0)} = 8.2292$$

Table 3　Chi-square test calculation table of the cultivated land that the average labor covered with 0 degree of mechanization

Number	Interval	n_i	p_i	np_i	$n_i - np_i$	$(n_i - np_i)^2/(np_i)$
1	$(-\infty, 1.11]$	9	0.0808	9.6960	−0.6960	0.0500
2	$(1.11, 1.21]$	14	0.1059	12.708	1.292	0.1314
3	$(1.21, 1.31]$	15	0.1653	19.836	−4.836	1.1790
4	$(1.31, 1.41]$	35	0.1997	23.964	11.036	5.0823
5	$(1.41, 1.51]$	19	0.1872	22.464	−3.464	0.5342
6	$(1.51, 1.61]$	17	0.136	16.32	0.68	0.0283
7	$(1.61, 1.71]$	6	0.0766	9.192	−3.192	1.1084
8	$(1.71, +\infty)$	5	0.0485	5.82	−0.82	0.1155
Σ	—	120	1	—	—	8.2292

Table 4　Chi-square test calculation table of the cultivated land that the average labor covered with 100% degree of mechanization

Number	Interval	n_i	p_i	np_i	$n_i - np_i$	$(n_i - np_i)^2/(np_i)$
1	$(-\infty, 7.22]$	7	0.0808	9.6960	−2.6960	0.7496
2	$(7.22, 7.52]$	11	0.1141	13.6920	−2.6920	0.5293
3	$(7.52, 7.82]$	26	0.1796	21.5520	4.4480	0.9180
4	$(7.82, 8.12]$	33	0.2126	25.5120	7.4880	2.1978
5	$(8.12, 8.42]$	26	0.1893	22.7160	3.2840	0.4748
6	$(8.42, 8.72]$	10	0.1285	15.4200	−5.4200	1.9051
7	$(8.72, +\infty)$	7	0.0951	11.4120	−4.4120	1.7057
Σ	—	120	1	—	—	8.4802

Degrees of freedom: $K-r-1=8-2-1=5$, if $\alpha=0.05$, the result showing below can be found through looking up the distribution table of χ^2.

$$\chi^2_\alpha(K-r-1) = x^2_{0.05}(5) = 11.071$$

Because of $\chi^2=8.2292<\chi^2_{0.05}(5)=11.071$, hypothesis H_0 is accepted. The data of the cultivated land that the average labor covered when the degree of mechanization is 0 can be recognized as obeying normal distribution. For the reason that:

$$\hat{\mu}(0) \pm 3\hat{\sigma}(0) = 1.3850 \pm 3 \times 0.1962$$
$$= [0.7965, 1.9735]$$

There are two abnormal data in the Table 1. They are two 2.0000. The abnormal data should be eliminated. The cultivated land that the average labor covered when the degree of mechanization is 0 without abnormal data is:

$$l(0) = \frac{1}{N_1'}\sum_{i=1}^{N_1'} l(0)_i = \frac{1}{118}\sum_{i=1}^{118} l(0)_i = 1.3746 \approx 1.37$$

When the degree of mechanization is 100%, the value of χ^2 is:

$$\chi^2 = \sum_{i=1}^{7} \frac{[n_i - np_i(100\%)]^2}{np_i(100\%)} = 8.4802$$

Degrees of freedom: $K-r-1=7-2-1=4$, if $\alpha=0.05$, the result showing below can be found through looking up the distribution table of χ^2.

$$\chi^2_\alpha(K-r-1) = x^2_{0.05}(4) = 9.488$$

Because, $\chi^2=8.4802<\chi^2_{0.05}(4)=9.488$, hypothesis H_0 is accepted. The data of the cultivated land that the average labor covered when the degree of mechanization is 100% can be recognized as obeying normal distribution. For the reason that:

$$\hat{\mu}(100\%) \pm 3\hat{\sigma}(100\%) = 7.9972 \pm 3 \times 0.5536$$
$$= [6.3364, 9.6580]$$

There are three abnormal data in the Table 2. They are 6.3333, 9.6667, 9.6667. The abnormal data should be eliminated. The cultivated land that the average labor covered when the degree of mechanization is 100% without abnormal data is:

$$l(100\%) = \frac{1}{N_1'}\sum_{i=1}^{N_1'} l(100\%)_i = \frac{1}{117}\sum_{i=1}^{117} l(100\%)_i$$
$$= 7.9829 \approx 7.98$$

Substituting $l(0) = 1.37$ hectares and $l(100\%) = 7.98$ hectares into Equation (9), Equation (5) can be obtained as follows:

$$l(x(t)) = \frac{10.93}{7.98 - 6x(t) + 1} \tag{23}$$

In addition, according to the statistical data of Suihua region, the data of the mechanization degree and the cultivated area in Suihua region from 2005 to 2014 is listed in Table 5.

Table 5　Mechanization degree and the cultivated area in Suihua region from 2005 to 2014

Year	Mechanization degree/%	Cultivated area/hm²
2005	46.41	1473333
2006	50.81	1469333
2007	54.23	1494000
2008	59.68	1615333
2009	64.96	1638000
2010	68.2	1806667
2011	73.96	1869333
2012	77.08	1919333
2013	82.06	1916667
2014	85.94	1906000

Make regression with the mechanization degree data from 2005 to 2014 in the Table 5.   The regression model can be established as:

$$\hat{x}(t) = \frac{1}{1 + 1.5226e^{-0.2104t}} \qquad (24)$$

In Equation (36), r=0.996 ($>r_{0.05}$) and F=500.3494 ($>F_{0.05}$). t is the time variable.

When t equals 1, 2, …,10, the corresponding years are the year of 2005, 2006, …, 2014.

The mechanization degree in Suihua region from 2015 to 2019 can be predicted based on the Equation (24).   To predict the cultivated area some interviews are made.   All the experts interviewed all agree that the cultivated area in Suihua region will not grow.   At the same time, the experts also say that the cultivated area will decrease in the future for the influences of expansion of the city, road construction and some other factors.   The decrease of cultivated area in Suihua region in the future can be counted as the average value of the cultivated area decreases in recent two years.   And the predicting results of mechanization degree and cultivated area are shown in the Table 6.

Table 6   Estimation results of mechanization degree and cultivated area

Year	2015	2016	2017	2018	2019
Mechanization degree/%	86.92	89.13	91.01	92.59	93.91
Cultivated area/hm^2	1899333	1892667	1886000	1879333	1872 667

Substituting Equation (23) into Equation (10), Equation (25) can be obtained as follows:

$$L(x(t)) = \frac{s(t)}{l(x(t))} = \frac{7.98 - 6.61x(t)}{10.93} \cdot s(t) \qquad (25)$$

Based on the Equation (25), the labor demand in planting industry in Suihua region from 2015 to 2019 can be estimated as follows.

Table 7   Estimation results of labor demand in planting industry in Suihua region from 2015 to 2019

Year	2015	2016	2017	2018	2019
Labor demand in planting industry (person)	388310	361651	338935	319780	303696

4   Conclusions

(1) Based on the actual production conditions, establishes an estimation model of labor demand in planting industry considering planting structure.   The model can accurately evaluate the quantity of labor demand in the planting industry, and make an estimation of the degree of mechanization and cultivated land in the future.   The quantity of labor demand in the future years will be forecasted at the same time.

(2) A case study of Suihua, Heilongjiang Province was conducted.   In the case study, the data of the cultivated land that the average labor covered were collected when the degrees of mechanization were 0 and 100% through survey.   Through the distribution of the survey data, a method of dealing with χ^2 distribution was applied and tested.   The results show that the data obtained by the questionnaire conform to the normal distribution.   On this basis, the abnormal data are eliminated and arranged to obtain the sample data of the cultivated land that the average labor covered area when the degrees of mechanization are 0 and 100%.   In addition, according to the data of the mechanization degree, and the measure of cultivated area, and the estimation model of labor demand for planting industry from 2005 to 2014, the demand of planting labor force in Suihua area were estimated to verify the accuracy and rationality of the model.

[References]

[1] Zhang L.   The research on methods of rural surplus labor estimates and prediction.   Harbin: Northeast Agricultural University, 2009. (in Chinese)

[2] Cang L.   Research on the degree of mechanization of plant industry and demand for labor.   Harbin: Northeast Agricultural University, 2009. (in Chinese)

[3] Kaushik B. A theory of surplus labor.   Tokyo: World Institute for Development Economics Research of the United Nations University, 1987.

[4] Ng Y C, Li S, Tsang S.   The incidence of surplus labor in rural China: A nonparametric estimation.   Journal of Comparative Economics, 2000; 28: 565–80.

[5] Ren J H, Zhang G S.   Analysis of the employment of rural surplus labor force in cities.   Agricultural Science

&Technology and Equipment, 2014; 5: 76–78. (in Chinese)

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

[7] Ou Y G, Malcolm W, Yang D T, Liu Q T, Zheng D K, Wang M M, et al. Mechanization technology: the key to sugarcane production in China. Int J Agric & Biol Eng, 2013; 6(1): 1–27.

[8] Lambeth R M. Scabbing the palouse: agricultural labor replacement and union busting in Southeast Washington. Labor History, 2015; 56(4): 1–20.

[9] Suo R X. Study on dynamic estimation method of rural surplus labor and its application. Harbin: Northeast Agricultural University, 2011. (in Chinese)

[10] Juhn C, Potter S. Changes in labor force participation in the United States. Journal of Economic Perspectives, 2006; 20(3): 27–46.

[11] Xie P X. Estimate for China's agriculture surplus labor. China Population: Resource and Environment. 2014; 14(1): 50–53. (in Chinese)

[12] 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.

[13] Cho S H, Chen Z, Poudyal N C. Spatial structure of agricultural production in China. Applied Economics, 2010; 42(16): 2031–2040.

[14] Suo R X, Wang F L, Sun X. 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)

[15] Li N. Study on the influence of the development of agricultural mechanization on the transformation of agricultural economic development mode in Xuchang of Henan. Agro Food Industry Hi-tech, 2017; 28(1): 270–274.

[16] Pradhan A, Jayasuriya H. Stataus and potentials of agricultural mechanization in sunsari district, NEPAL.

Applied Engineering in Agriculture, 2016; 32(6): 759–768.

[17] 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.

[18] Wang F L, Suo R X, Zhang L, Cang L. Model of relationships between planting industry mechanization degree and labor demand. Transactions of the CSAE, 2010; 26(9): 181–184. (in Chinese)

[19] Yu Y T. Agricultural Mechanization Engineering. Beijing: China Outlook Press, 1987.

[20] Dai Y Z, Tuo B Z, Lin W, Ouyang H. A prevision of the trend of equipping Shuangfeng county with the farm machinery by 'Delphi Method'. Journal of Northeast Agricultural University, 1983; 4: 57–65. (in Chinese)

[21] Zhang L. Experience in teaching methods of probability theory and mathematical statistics. Agro Food Industry Hi-Tech, 2017; 28(1): 963–967.

[22] Kosareva N, Krylovas A. A numerical experiment on mathematical model of forecasting the results of knowledge testing. Technological and Economic Development of Economy, 2017; 17(1): 42–61.

[23] Bell R C, Glen A G. Experiences teaching probability and statistics with personal laptops in the classroom daily. American Statistician, 2008; 62(2): 155–160.

[24] Wang F L, Shi G H, Wang J Q. Study on the method of getting and dealing with some technical and economic indexes of tractor. Transactions of the CSAM, 2006; 37(6): 76–78. (in Chinese)

[25] Wu M C. Basic principles and methods of mathematical statistics. Chengdu: Northwest University of Finance and Economics Press, 2006. (in Chinese)

[26] Rice J A. Mathematical statistics and data analysis. USA: Duxbury Resource Center, 2006.

[27] Lutz R, Musio M. An alternative mathematical foundation for statistics. Acta Applicandae Mathematicae, 2005; 89(1): 217–249.