*J Integral Sci*, 2018, Volume 1, Issue 1, 12-18. ISSN: 2581-5679

www.jisciences.com

Research Article

**Journal of Integral Sciences**

# Statistical Analysis of Factors Affecting the Production of Wheat Crop In Gulomakeda Woreda, Eastern Tigray Zone, Ethiopia.

Venkata Subba Rao U [1,*], Rahiman SA [2]

[1] Department of Statistics, College of Natural& Computational Sciences, Adigrat University; Ethiopia.

[2] Department of Environmental Sciences, College of Agriculture & Environmental Sciences, Adigrat University, Ethiopia.

* Corresponding Author: E-mail: sr.uppu@gmail.com, Mobile: +251 914439631.

Received: 19th March 2018; Revised: 28th March 2018 Accepted: 30th March 2018

**Abstract**

The present research is done based on the major factors which are affecting the production of wheat crop in Gulomekada woreda, Eastern Tigray, Ethiopia. The Gulomekada is one of the main wheat producing area in the tigray region. But, recently the production of wheat decreased due to different reasons such as land size, pesticide, fertilizer, temperature and rainfall. The present study was about the statistical analysis of the factors affecting the production of wheat crop. The collected data was a secondary source of data about the factors affecting the production of wheat and the data analyzed by the descriptive and inferential statistics. From the results of the present study, the average production of wheat in the Gulomakeda woreda for the past 10 years was 14719.25 quintal per hectare. The average land size, pesticide, fertilizer, temperature and rainfall in the past 10 years 2476.68hectare, 1755.00lts, 4054.70qu, 20.50℃ and 933.90mm respectively. In addition, from the multiple linear regression models some independent variables (factors) like land size, fertilizer, temperature and rainfall are statistically significant at 5% level of significance.

**Key words:** Wheat, Gulomakeda, Woreda, Tigray and Ethiopia.

## 1. Introduction

Wheat is the third cultivating cereal species in the world after maize and rice. This grain is grown on more land area than any other commercial food. Wheat production has been increasing steadily over the years contributing to insufficient food across the world that are presently being filled with import and food aid from developing counties. Many parts of the sub-Saharan Africa region are biophysically suitable for wheat production. Ethiopia is second largest producer of wheat in sub-Sahara region. Wheat is one of the major cereal crops grown in the Ethiopian highlands. At present, wheat is produced solely under rain-fed conditions (Tesfaye *et al.,* 1991; Cimmyt, 2000; CSA, 2011), but currently the size of land under cultivation is shirking in Ethiopia due to different reasons. The above situation is true in the case of Gulomakeda woreda, eastern Tigray zone, Ethiopia. Administratively this woreda is structured into 18 rural kebeles and has total area of 1596.12 square kilometers. Agriculture is the main income source of the woreda and wheat production is the highest portion of all other crops. If the production of wheat shrinking the economic status of the people in this region will effect, in this point of view, the study tries to assess with identification of over all factors of wheat crop production in this woreda.

## 2. Materials and Methods

### *2.1 Study area*

Gulomakeda is one of the woredas in the Tigray region of Ethiopia, with an area of 1,596.12 square kilometers, Gulomakeda has a population density of 52.78, which is less than the Zone average of 56.93 persons per square kilometer. A total of 18,365 households were counted in this woreda, resulting in an average of 4.59 persons to a household, and 17,673 housing units. The maximum people of this woreda are farmers, who held an average of 0.37 hectares of land. The land under cultivation in this woreda is 65.36%. In planted cereals, 3.22% in pulses, and 0.85% in oilseeds (Reports on area and production, 2007).

*2.2 Method of Data Collection*

In this study, the data was collected by using secondary source of data. This will be collected by woreda Glumekada Agricultural and rural development office, of the year 1996-2005 Ethiopian Calendar (E.C) (Appendix 1) (Cochran, 1997).

*2.3 Method of Data Analysis/Statistical Analysis*

*2.3.1 Variable Identification/Variable Considered In the Study*

The dependent and independent variables were studied in the present research.

Dependent (response) variables: Yield of wheat crop(Y) (quintal)

Independent (factor or explanatory) variables:

A. Temperature($X_1$)(℃)
B. Rain fall($X_2$)(mm)
C. Pesticides($X_3$)(Lt)
D. Fertilizer uses($X_4$)(quintal)
E. Land size($X_5$)(hectare)

*2.3.2 Methods of statistical analysis*

To accomplish the data, the two broad areas of statistics which are descriptive and inferential statistics were used (Bryman Alan and Cramer Duncan, 2011; Farrar Donald and Glauber Robert, 1967; Krishna Kumar, 1975).

*2.4 Multiple linear regression models*

The primary objective of regression is to develop a regression model, to explain the relation between one or more variables in a given population. A particular form of regression model depends up on the nature of the problem under study and the type of data variables. Multiple linear regressions contain two or more independent variables and one dependent variable.

The general form of a multiple linear regression model is given by

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + - - - + \beta_k X_{ki} + \varepsilon_{i i=1, 2, \ldots, n}$$

Where $\beta_0$ is the intercept and $\beta_1, \beta_2 \ldots \beta_k$ is coefficient of the variable $X_1, X_2 \ldots$

But in our research paper we use five independent variables and one dependent variable. That means

$$Y=\beta_0+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3+\beta_4 X_4+\beta_5 X_5+\varepsilon_i$$

Where Y=yield of wheat crop; $X_1$=temperature in (℃);$X_2$=rainfall in (mm);$X_3$=pesticides in (Lt);$X_4$=fertilizer in (quintal);$X_5$=land in (hectare);$\varepsilon_i$ =the random error term

The above model shows that:

- Constant (βo) is the value of wheat production when the all independent variables($X_1$, $X_2$, $X_3$, $X_4$ and $X_5$) are zero.
- $\beta_1$ is the change in total production of wheat when the temperature($X_1$) increases by one, keeping rainfall($X_2$), pesticides($X_3$), fertilizer($X_4$) and land size ($X_5$) constant.
- $\beta_2$is the change in total production when $X_2$ increase by on unit that is the production increase by $\beta_2$ unit keeping $X_1$, $X_3$, $X_4$ and $X_5$ constant.
- $\beta_3$ is the change in total production when $X_3$ increase by one unit ,the production increase by $\beta_3$ units keeping $X_1$,$X_2$,$X_4$ and $X_5$ constant
- $\beta_4$ is the change in total production when $X_4$ increase by one unit ,the production increase by $\beta_4$ units keeping $X_1$,$X_2$,$X_3$ and $X_5$ constant
- $\beta_5$ is the change in total production when $X_5$ increase by one unit ,the production increase by $\beta_5$ units keeping $X_1$,$X_2$,$X_3$ and $X_4$ constant

*2.5 Hypothesis testing for significance of regression*

Ho: $\beta_1=\beta_2=---\beta_5= 0$

$H_1$:$\beta i\neq 0$, for at least one i

The total sum squares is given by: SST=SSR+SSE

Where SSR=sum square of regression

SSE=sum square of residual (errors)

To test Ho: $\beta_1=\beta_2---=\beta_5$, first we compute F calculated and F tabulation as follow.

$$MSR=\frac{SSR}{K}$$

$$MSE=\frac{SSE}{n-k-1}$$

$$F_{cal}=\frac{MSR}{MSE}$$

Where, MSR is mean square regression.

MSE is mean square residual

We compare $F_{cal}$ with Fά (k, n-k-1) and we reject Ho If $F_{cal}$>Fά (k, n-k-1).

Where 'k' represents number of independent variables and 'n' is number of observation

If Ho is rejected it means that the contribution of xi in estimating Y is significant.

If Ho is accepted it means that the contribution of xi in estimating Y is not significant.

Of individual regression coefficients

Null hypothesis: Ho: βi=0 individual factors have no effect

Alternative hypothesis: $H_1$: βi≠0 i=1, 2…5.

Test statistic ($t_{test}$) $= \frac{\beta i}{se(\beta i)}$

Decision rule: reject Ho if $t_{cal} > t\, \alpha/2(n-1)$

If Ho is rejected, it means that the contribution of $x_i$ in estimating

Y is significant.

If Ho is accepted, it means that the contribution of $x_i$ in estimating

Y is not significant.

- t- Distribution used to test the significance of one parameter. When the numbers of observation are less than 30 and the distribution of the population is normally distributed.
- Use of F-test: To test for the significance of the overall model, before considering the significance of individual variables.
- Use of t-test: It is used to determine if the individual coefficient for each independent variable represents a significant contribution to the overall model.

## 4. Results and Discussion

*4.1 Descriptive Analysis*

The main objective of this study was to identify impacts that affect the yield of wheat crop in Gulomakeda woreda. The secondary data was collected on the yield of wheat crop issues during the period of 1996 up to 2005 E.C. The descriptive statistics of SPSS output with 10 years of wheat crop production are given in the Table 1.

The maximum yield of wheat crop is 15483qu which is recorded in 2005; at the year of maximum crop yield recorded the amount of cultivated land size, pesticide, fertilizer, temperature and rainfall was 2527hectare, 2504Lt, 7865qu, 24℃ and 1150mm respectively.

The minimum yield of wheat crop was14054qu which was recorded in 1999, at the year when the minimum wheat crop production recorded the amount of cultivated land size, pesticide, fertilizer, temperature and rainfall was 2412hectare,1120Lt, 2263qu ,17℃ and 800mm respectively.

**Table 1.** Descriptive Statistics of wheat production at Gulomakeda woreda

| Descriptive Statistics | | | | | |
|---|---|---|---|---|---|
| Name of the parameter | N | Minimum | Maximum | Mean | Std. Deviation |
| Yield (qu) | 10 | 14054 | 15483 | 14719.25 | 473.21 |
| Land size (hr) | 10 | 2412 | 2527 | 2476.68 | 34.25 |
| Pesticide (lts) | 10 | 1120 | 2504 | 1755.00 | 586.16 |
| Fertilizer (qu) | 10 | 2263 | 7865 | 4054.70 | 1842.96 |
| Temperature (℃) | 10 | 17 | 24 | 20.50 | 2.59 |
| Rainfall (mm) | 10 | 800 | 1150 | 933.90 | 129.13 |

**Table 2.** Model Summary

| Model Summary [b] | | | | | |
|---|---|---|---|---|---|
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin-Watson |
| 1 | .990[a] | .981 | .957 | 98.639 | 2.257 |

a. Predictors: (Constant), rainfall, temperature, pesticide, land size, fertilizer

**Tab1e 3.** ANOVA analysis of Independent factors on dependent variable

| ANOVA [a] | | | | | |
|---|---|---|---|---|---|
| Model | Sum of Squares | df | Mean Square | F | Sig. |
| Regression | 1976473.303 | 5 | 395294.661 | 40.628 | .002[b] |
| Residual | 38918.218 | 4 | 9729.555 | | |
| Total | 2015391.521 | 9 | | | |

a. Dependent Variable: yield

b. Predictors: (Constant), rainfall, temperature, pesticide, land size, fertilizer

b. Model Summary

The mean values of cultivated land size, pesticide, fertilizer, temperature and rainfall are 2476.68hectare, 1755.00Lt, 4054.70qu, 20.50℃and 933.90mm respectively which shows that land size is the most used up factor.

From the output the standard deviation of production of wheat in quintal, cultivated land size in hectare, pesticide in liter, fertilizer in quintal, temperature Celsius and rainfall millimeter are 473.215qu, 34.255hectare, 586.168Lt, 1842.963qu, 2.593℃and 129.135mm respectively (Barlow Jesse, 1993).

*3.2 Multiple linear regression analysis*

The statistical model that was used in the study is multiple linear regression models. The general model for multiple linear regression analysis is used to check the effects of many quantitative independent variables on single response (Pedhazur Elazar, 1982). Mathematically, the model is given as:

$$Y = \beta_o + \beta_1 X_1 + \beta_2 X_2 + ....... + \beta_k X_k + \varepsilon_{ij}$$

Before we use the results and interpret, we have to check model adequacy and whether assumptions are satisfied or not. The model summary reports the strength of the relationship between the model and  the dependent variable (Table 2). According to the given fitted model, the value of R = 99% this indicates that there is strong relationship between variables and also the value of $R^2$ = 98.1% of the variation in the production of wheat crop is explained by the five explanatory variables; land, pesticides, fertilizer, temperature and rainfall are jointly and linearly. Hence the model is adequate.

*3.3 Hypothesis testing for the model*

This method is used to test the effect of the independent variables on the dependent variable.

ANOVA is a useful test of model's ability to explain any variation in the dependent variable; it does not directly address the strength of that relationship (Table 3).

*3.3.1 Test for significance of regression*

The objectives are to determine if there is a linear relationship between the response variable and any one of the regress or independent variables. The hypotheses are as follow.

Step 1; Ho: $\beta_1=\beta_2=\beta_3= \beta_4= \beta_5=0$VS $H_1$: $\beta_i\neq0$, for at least one **i** is defer

Step 2; $\alpha$=0.05

Step 3; Test statistics, $F_{cal}$=40.628

Step4; P-value =.002

Step5; Decision; since p-value=.002is <$\alpha$-value= 0.05. we reject Ho. Therefore, we can conclude that the overall regression model is statistically significant, that means at least one of the parameters or coefficients of explanatory variables are different from zero.

*3.4 Interpretation of the model coefficient*

Production=30533.872-6.102land+0.038pesticide + 0.401fertilizer -48.570temprature -1.498rainfall.

The intercept $\beta_0$=30533.872 indicates when the value of cultivated land size, pesticide, fertilizer, temperature and rainfall are assume zero the value of production of wheat, is 30533.872.

Firstly, the slope of land size $\beta_1$=-6.102 is negative indicating that there is inverse (negative) linear relationship between land size and the production of wheat crop.

The amount of change in the production of wheat crop when pesticide, fertilizer, temperature and rainfall change by one unit is 0.038, 0.401, -48.570 and -1.498respectively assuming that the production of wheat crop remains constant.

Secondly, the slope of fertilizer $\beta_3$=0.401 is positive indicating that there is direct (positive) relationship between fertilizer and the production of wheat crop. The amount of change in the production of wheat crop is when the land size, pesticide, temperature and rainfall change by one unit is- 6.102, 0.038, -48.570and -1.498respectively assuming that the production of wheat crop remains constant.

Thirdly, the slope of temperature $\beta_4$= -48.570 is negative indicating that there is inverse (negative) linear relationship between temperature and the production of wheat crop. The amount of change in the production of wheat crop is when the land size, pesticide, fertilizer and rainfall change by one unit is - 6.102, 0.038, 0.401and -1.498

respectively assuming that the production of wheat crop remains constant.

Fourthly, the slope of rainfall $\beta_5$= -1.498 is negative indicating that there is inverse (negative) relationship between rainfall and the production of wheat crop.

The amount of change in the production of wheat crop when the land size, pesticides, fertilizer and temperature change by one unit is -6.102, 0.038, 0.401and -48.570 respectively assuming that the production of wheat crop remains constant (Table 4).

**Tab1e 4.** Result of regression coefficient

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | Co linearity Statistics | |
|---|---|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | | Tolerance | VIF |
| 1 | (Constant) | 30533.872 | 3687.688 | | 8.280 | .001 | | |
| | land size | -6.102 | 1.554 | -.442 | -3.926 | .017 | .381 | 2.622 |
| | Pesticide | .038 | .088 | .047 | .436 | .685 | .407 | 2.456 |
| | Fertilizer | .401 | .046 | 1.562 | 8.720 | .001 | .150 | 6.650 |
| | temperature | -48.570 | 15.631 | -.266 | -3.107 | .036 | .658 | 1.519 |
| | Rainfall | -1.498 | .480 | -.409 | -3.124 | .035 | .282 | 3.547 |

a. Dependent Variable: yield

*3.5 Testing individual parameters*

Since we have rejected the null hypothesis we have to test which variable is responsible for the rejection. From table of multiple regressions for the model, land size, fertilizer, temperature and rainfall has a significance effect both individually and in group on the production. Because their p - values are less than 5% significance level value.

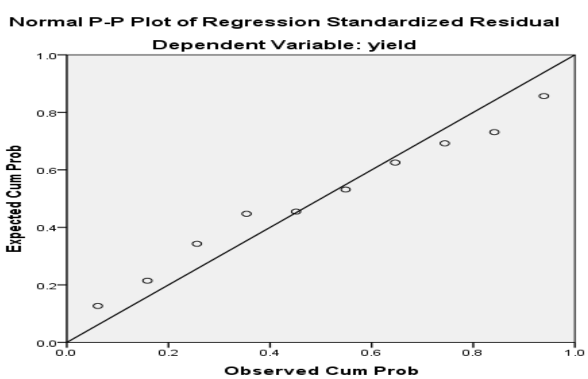


**Fig. 1.** normal-p plot regression standardized residual

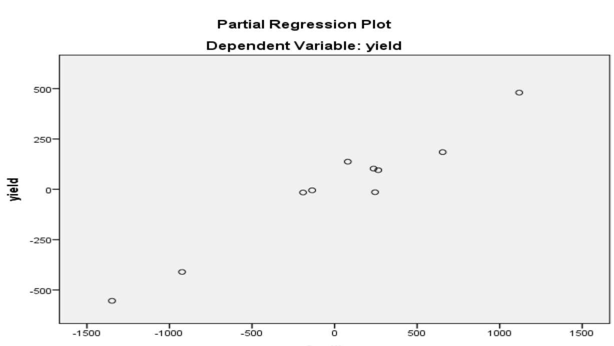


**Fig 2.** Normal probability plots of residuals

*3.6 Test of Model Diagnosis*

*3.6.1 Linearity*

The Fig 1 is approximately normal it show that there is approximately a linear relationship between response and explanatory variables.

*3.6.2 Homoscedicity*

Since the points, in the Fig 2 indicating a sharp downward and upward curve at both extremis indicating the tail of distribution is to be considered normal. Points in the plot are not dispersed at random fashion; this indicates that the assumption of Homoscedasticity is hold. This means the models are well defined.

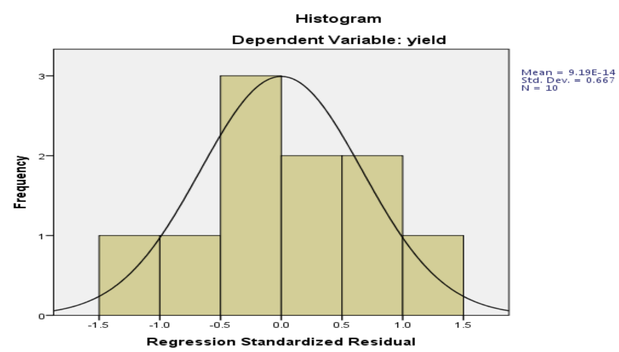


**Fig 3.** Histogram plot

*3.6.3 Normality of Error Term*

The below Fig 3 shows, that the error terms are normally distributed and approximately with mean zero and with constant variance. This implies that independent variable X and error terms are independent. This means the models are well defined.

*3.6.4 Autocorrelation*

The above Table 2 shows that the value of Durbin –Watson test 2.257 is positive this indicates that there is a positive autocorrelation .This implies that our model is well defined.

*3.6.5 Multi-colinerity*

The co linearity statistics tells as whether multicolinarity exists in the model or not. From the above result, we see that the variance inflation factor (VIF) is less than 10; hence, we can conclude that there is no problem of multicolinerity presents in the data. Of course, Tolerance of all variables is approximately closer to one, which implies that there is a little multicolinerity in the model**.**

*3.7 Testing the Assumptions of Multiple Linear Regression*

*Normality assumption*: Distributed with mean zero and variance σ2 can be tested by plotting residual against the cumulative the error terms are normally probability. The above Fig 1 is approximately normal it show that there is approximately a linear relationship between response and explanatory variables and which indicates our model (Y=β0+β1X1+……+β5X5+€i) is adequate. Therefore the linearity assumption is valid. Assumption of Homoscedasticity: Error terms have constant variance.

Since the points, in the Fig 2 indicating a sharp downward and upward curve at both extremis indicating the tail of distribution is to be considered normal. Points in the plot are not dispersed at random fashion; this indicates that the assumption of Homoscedasticity is hold. This means the models are well defined.

The above Fig 3 shows that the error terms are normally distributed approximately with mean zero and constant variance. This implies that independent variable X and error terms are independent. This means the models are well defined (Chatterjee and Hadi, 2000). The shape of the histogram should approximately follow the shape of the normal curve. This histogram is acceptably close to normal curve. Therefore, the assumption of normality holds.

## 4. Conclusion

Depending on the analysis of the result and discussion, we conclude the following main points.

The average production of wheat in the Gulomakeda woreda for the past 10 years was 14719.25 quintal per hectare. The average land size, pesticide, fertilizer, temperature and rainfall in the past 10 years 2476.68hectare, 1755.00lts, 4054.70qu, 20.50℃ and 933.90mm respectively.

In addition, from the multiple linear regression models some independent variables (factors) like land size, fertilizer, temperature and rainfall are statistically significant at 5% level of significance.

Furthermore, from the multiple linear regression models the result indicates that the relationships between the amount of wheat production and the factors (land size, temperature, and rainfall) are negative linear relationships (impact); however, pesticide and fertilizer is positive linear relationship.

In Gulomakeda woreda, production of wheat crop is decreasing, so that in order to improve the production and to ensure the socio-economic well-being of residents of were dagulomakeda the following measure should be taken. The farmer should have to use any kind of improvement such as enough improved seeds, modern farming techniques and etc. The agricultural sectors should be highly encourages and lead the farmers by giving more technical (professional) follow up, to get highly qualified yields and to improve the production technique. Since the production of wheat crop depends on land size, fertilizer, temperature and rainfall the metrological agency and agricultural office should work together with each other on this issue.

Since there is no data that collected for long period of time, as result we cannot study the effect of production in detail and we cannot know the normality of the data. Therefore, the agricultural office workers and each employment should be accomplish their responsibility at right time and right place in recording data from time to time for the future researchers.

**Acknowledgments**

The authors are thankful to Agricultural and rural development officers, Gulomakeda wereda, Tigray, Ethiopia for their support in the data collection.

**Conflict of Interest**

Author has none to declare.

**References**

Report on Area and Production - Tigray Region. 2007. Central Statistical Authority of Ethiopia. Agricultural Sample Survey (AgSE2001).

CIMMYT. 2000. The Eleventh Regional Wheat Workshop for Eastern. Central and Southern Africa. Addis Ababa, Ethiopia.

CSA. April, 2011. Report on Area and Production of Major crops. Ethiopian Agricultural Sample Survey Private Peasant Holdings, Meher Season (2010/11 (2003 E.C.) – Volume I. Statistical Bulletin. Addis Ababa: Central Statistical Agency.

Tesfaye T, Getachew B, Worede M. 1991. Morphological diversity in tetraploid wheat landrace populations from the central highlands of Ethiopia. *Hereditar*, 114, 171-176.

Cochran WG. 1997. Sampling techniques.3rd edition, john Wiley and sons, inc…New York".

Bryman Alan, Cramer Duncan. 2011. Quantitative Data Analysis with IBM SPSS 17, 18 and 19: A Guide for Social Scientists. New York.

Farrar Donald E, Glauber Robert R. 1967. Multicollinearity in Regression Analysis: The Problem Revisited. *Review of Economics and Statistics,* 49 (1), 92–107.

Krishna Kumar T. 1975. Multicollinearity in Regression Analysis. *Review of Economics and Statistics,* 57 (3), 365–366.

Barlow Jesse L. 1993. Chapter 9: Numerical aspects of Solving Linear Least Squares Problems. In Rao, C.R. Computational Statistics. Handbook of Statistics 9. North-Holland.

Pedhazur Elazar J. 1982. Multiple regression in behavioral research: Explanation and prediction (2nd ed.). New York.

Chatterjee S, Hadi AS, Price B. 2000. Regression Analysis by Example (Third ed.). John Wiley and Sons.

**Appendix:**

Raw data of yield of wheat crop production of Gulomakeda woreda (From 1996 – 2005 EC).

| Year | Yield (quintal) | land size(hectare) | pesticide (Lt) | fertilizer (quintal) | Temperature (°C) | rainfall(mm) |
|---|---|---|---|---|---|---|
| 1996 | 14896 | 2475 | 1254 | 3625 | 17 | 800 |
| 1997 | 14816 | 2466 | 2135 | 3987 | 24 | 812 |
| 1998 | 15296 | 2527 | 1201 | 6897 | 20 | 1150 |
| 1999 | 14054 | 2494 | 2406 | 3201 | 24 | 956 |
| 2000 | 14274 | 2499 | 2504 | 2965 | 23 | 821 |
| 2001 | 14367 | 2459 | 2412 | 2564 | 18 | 863 |
| 2002 | 14239 | 2452 | 1265 | 2263 | 20 | 869 |
| 2003 | 14816 | 2463 | 2012 | 3620 | 18 | 968 |
| 2004 | 14952 | 2412 | 1241 | 3560 | 19 | 950 |
| 2005 | 15483 | 2520 | 1120 | 7865 | 22 | 1150 |