**Journal of the Institute of Engineering**
Volume 18, No. 1
ISSN: 1810-3383
Published: February 2025

Printed in Nepal

# International Air Passenger Demand Modeling in Nepal

Nishub Luitel [a], Pradeep Kumar Shrestha [b], Ashok Kumar Shrestha [c]

[a, b, c] *Department of Civil Engineering, Pulchowk Campus,IOE, Tribhuvan University, Nepal*
**Corresponding Author**: [a]luitelnishub@gmail.com

Received: 2023-10-29 Revised: 2023-12-06 Accepted: 2023-12-21

**Abstract:**
Air travel demand forecasting is an important task for the concerned civil aviation authority as well as airlines company. This paper intends to analyze and forecast international air travel demand in Nepal. Econometric variables like GDP, CPI, remittance, employment migration, tourist arrivals, exchange rates, purchasing power parity etc were taken as an explanatory variable for the demand generated in aviation industry. A Multiple linear regression model was developed where above mentioned variables were used as explanatory variable and the statistical result showed almost perfect correlation between themselves which made the model biased in terms of coefficients. Furthermore, several models considering different combinations of independent variables were developed. Finally a regression model where exchange rate, number of labour permit, number of tourist arrival taken as explanatory variable was developed. The developed model showed satisfactory result and was taken as demand model for the forecast.

**Keywords**: Demand modeling, Time series data, Forecasting

## 1. Introduction

Transportation is the backbone for the development of any nation. Historical evidences show that the countries having well developed transportation has developed faster. Due to the worldwide globalization, aviation industry has grown much faster than any other mode of transportation. In such growing market, proper planning of infrastructure is required and hence forecasting the passenger demand for the development of facilities like airport terminal building. Runways, aprons etc is required.

Forecasts are informed prediction of future aviation activities that are supported by careful assessment and analysis. It serves as basis for effective decision about the terminal plan. Because of the impacts of air travel on the transportation networks and on the environment, forecasts of the future demand for air travel and knowledge of its determining factors are essential components in the formation of transportation policies.

The increasing air travel demand has led to a scenario where the capacity of Tribhuwan International Airport (TIA) to handle air traffic has almost reached its saturation level. Indeed, the changing of demand between airports not only has an effect on the commercial feasibility of the single airports, but can have significant effects on the support structure of the airports (auxiliary businesses), as well as on seemingly less related businesses (e.g. local hotels). One of the most important consequences of increase in demand is the need for improvements in surface access facilities; this can cause major problems in metropolitan areas (in our case Kathmandu metropolitan city) that often already suffer from severe ground-transport congestion. The effects of decrease in demand are equally significant, as the survival of many local businesses and related jobs (e.g. taxi companies) depends to a large degree on the success of the airport that was often responsible for their creation in the first place. Hence a need for forecasting air travel demand is of utmost importance.

## 2. Previous Works

Numerous studies have been carried out on the topic demand forecasting. The literatures required for the model development were reviewed thoroughly and are described below.

Wang and Song [1] provided a brief review of the world air travel demand between 1950 and 2008. This study presented literature survey on air travel demand studies and provides a comprehensive review of the existing studies. Rengaraju and Arasan [2] has done extensive studies where they calibrated a city pair model for the demand for domestic air travel based on the air travel in 40 city pairs in india. Karlaftis et al. [3] established a time series model with econometric independent descriptive variables to forecast the international passengers at Frankfurt airport. The world value of Gross Domesctic Product ( GDP) per capita was found to be the finest explanatory variable for international passengers. Erraitab [4] developed a regression model for demand forecast in morocco. The primary objective of this paper was to formulate an econometric model for the examination, evaluation, and projection of air travel demand in Morocco. The procedure of conducting an all-possible regression was employed to select the pertinent variables. Among the various models tested, the one incorporating the consumer price index, gross national product, household final consumption per capita, and international tourist arrivals emerged as the most suitable model for representing air travel demand in Morocco. Seraj et al. [5] developed several models for the air travel demand with different combinations of explanatory variables utilizing stepwise regression technique. The paper provided an in-depth account of the procedures undertaken to create econometric models for forecasting international air travel demand within Saudi Arabia. Efforts were made to employ econometric models to examine and predict air traffic by establishing statistical connections between specific factors that influence demand and the corresponding traffic levels. Based on the statistical assessments used to evaluate these models, it was determined that the primary determinants of international air travel in Saudi Arabia were the population size and total expenditure. Bastola [6] developed an econometric model to estimate demand in passenger air transport for Nepal. The model assumed that passengers carried by Nepali air transport depend on the GDP and number of tourist arrival. Dargay and Hanly [7] developed a model to examine the demand for air travel to and from the United Kingdom (UK) over the past decade and to explore the factors that influenced its growth. The factors under consideration included income, airfares, foreign trade, exchange rates, and domestic price levels. Dynamic econometric models for air travel were estimated using pooled time-series cross-section data.

## 3. Methodology

### 3.1 Data Collection

Only secondary sources of data was used for modelling which was extracted from the reports published by several institutions. Passenger movement data of TIA from 1998 to 2019 was collected from the reports published by CAAN. Similarly, Econometric variables like GDP, CPI, exchange rates, PPP, required for the analysis have been extracted from the reports published by the world bank.

### 3.2 Identification of Variables

In forecasting air-travel demand, it is of critical importance both to ensure the statistical validity of the models, and to select the models with the higher potential forecasting accuracy. In our study several variables that might affect the air travel demand has been considered and analysis is done to select the best model which is statistically valid and have a good forecasting power. Following variables are taken as independent and dependent variable for demand forecasting model and are described briefly below.

**International Air passenger demand**:It is the total number of passenger movement at TIA international terminal.The volume of passengers recorded annually is the dependent variable in this study. Required data was collected from the reports published by Civil Aviation Authority of Nepal (CAAN)[8].Data from 1998 to 2019 was included in modeling.

**GDP (gross domestic product)**:It is the total monetary or market value of all the finished goods and services produced within a country's borders in a specific time period.The GDP is the total of all value added created in an economy. Historic data of GDP for Nepal was extracted from the database of the world bank[9] and measurement unit was inUnited States dollars (USD).

**CPI(consumer price index):** The Consumer Price Index (CPI) is a measure that examines the weighted average of prices of a basket of consumer goods and services, such as transportation, food, and medical care. CPI is calculated by taking the cost of the basket in a given year and dividing it by the cost of the same basket in the base year, then multiplying by 100. The changes in the CPI were used to assess price changes associated with the cost of living. 2010 was considered base year[9] and CPI being ratio was unit less.

**Labour permit:** It gives the number of Nepali Citizen who has been given authority to work in foreign countries. Required data was collected from the report published by Ministry of labour, Employment and Social Security (MOLESS)[10]. The unit for measurement was number of person.

**Remittance:** Remittances are funds transferred from migrants to their home country. They are the private savings of workers and families that are spent in the home country for food, clothing and other expenditures, and which drive the home economy. Data for remittance was extracted from the database of the World Bank[9]. The unit for measurement was in USD

**Number of tourist arrivals:** It is the total number of foreign citizens who has arrived to Nepal as a tourist. Required data was collected from the report published by Ministry of Culture, Tourism and Civil aviation( MOCTCA)[11]. The unit for measurement was number of person.

**Purchasing power parity (PPP) conversion factor:** it is a spatial price deflator and currency converter that controls for price level differences between countries. PPP being ratio was unit less[9].

**Official exchange rate:** it refers to the actual, principal exchange rate and is an annual average based on monthly averages (local currency units relative to U.S. dollars). Required data was collected from the database published by the World Bank[9]. The unit for measurement was in USD

### 3.3 Model Formulation

Regression analysis was used to develop the relationship between dependent variable and one independent or more independent variable. Multiple linear regression is performed when there is more than one independent or explanatory variables. Mathematical representation of multiple regression is in the form of:

$$y = a + b \cdot x_1 + c \cdot x_2 + \cdots \tag{1}$$

where,

$y$ = Dependent variable

$x_1, x_2, \ldots$ = Independent variables

$a$ = Intercept

$b, c, \ldots$ = slope

In multiple linear regression (MLR), the coefficients are estimated using a technique called Ordinary Least Squares (OLS). OLS aims to find the coefficients that minimize the sum of the squared differences between the observed values of the dependent variable and the values predicted by the regression model. The goal is to find the values of a, b, c, ..., that minimize the sum of squared errors (SSE), which is the sum of the squared differences between the observed Y values and the predicted Y values. The OLS estimation method minimizes the following cost function:

$$\mathrm{SSE} = \sum_i (Y_i - \hat{Y}_i)^2 \tag{2}$$

Where,

$Y_i$ is the observed value of the dependent variable.

$\hat{Y}_i$ is the predicted value of the dependent variable based on the model.

### 3.4 Criteria for Model Selection

The Coefficient of determination ( $R^2$ ) is used to estimate goodness of fit of the function to the data. The $R^2$ value represents the fraction of the overall variance of the 'dependent' variable that is explained by the 'independent' variable. Multiple linear regression is built on several key assumptions, and validating these assumptions is crucial to ensure the reliability and accuracy of the model. Some important assumptions of multiple linear regression and methods for their validation is given below.

**Linearity Assumption:**

Assumption: The relationship between the dependent variable and the independent variables is linear. This means that the change in the dependent variable is proportional to the change in each independent variable.

Validation: Creating scatter plots of the dependent variable against each independent variable helps to visually check for linearity.

**Independence Assumption:**

Assumption: The residuals are independent of each other. This means that the error in predicting one observation does not depend on the errors in predicting other observations.

Validation: Durbin-Watson statistical test that checks for autocorrelation in the residuals. Durbin Watson significance table can be used to validate the assumption.

**Homoscedasticity Assumption:**

Assumption: The variance of the residuals is constant across all levels of the independent variables. In other words, the spread of residuals is the same for all values of the independent variables.

Validation: Plot residuals against the predicted values or independent variables can visually assess homoscedasticity. Homoscedasticity assumption is supposed to be met if there is consistent spread of residuals.

**Normality of Residuals:**

Assumption: The residuals follow a normal distribution.

Validation: Examining Normal probability plot (Q-Q plot) whether the residuals follow a straight line on a Q-Q plot can visually check Normality.

**No Perfect Multicollinearity Assumption:**

Assumption: There is no perfect linear relationship among the independent variables. In other words, no independent variable can be exactly predicted by a linear combination of the others.

Validation: Calculate the Variance Inflation Factor (VIF) for each independent variable. A high VIF (typically > 10) suggests multicollinearity.

## 4. Results and Discussion

A regression model was developed where international air passenger demand at (TIA) was the dependent variable and PPP conversion factor, number of labour permit, Tourist arrival, Exchange rate, world GDP, GDP per capita, Remittance, CPI, GDP constant Local currency unit ( LCU) was taken as independent variable. Model summary given in Table 1 revealed $R^2$ value of 0.998 which indicates that the model accounts for 99.8 % variability observed in the data. From Table 3, it was found that all variables except tourist arrival (p=0.000) and number of labour permit(p=0.05) and exchange rate(p=0.022) was found insignificant. Besides coefficients of the variables being insignificant, multicollinearity, as indicated by VIF value was found among the predictor variables. Hence, the model including above mentioned variable was found to be statistically unjustified. (i.e. p>0.05). As a finding from result of first regression, variables highly correlated as shown in Table 4 and variables having similar meaning were excluded from the analysis.

**Table 1:** Model summary-Model 1

| *R* | $R^2$ | Adjusted $R^2$ | Std. error | D watson |
|---|---|---|---|---|
| 0.999 | 0.998 | 0.996 | 0.70 | 2.48 |

**Table 2:** ANOVA-Model 1

| model | S S | DF | Mean sqr | F | sig |
|---|---|---|---|---|---|
| Regression | 29.081 | 9 | 3.231 | 642.88 | 0.00 |
| Residual | 0.060 | 12 | 0.05 | | |
| Total | 29.14 | 21 | | | |

**Table 3:** Model Coefficients-Model 1

| Model | unstandard coef | standard coef | t | sig | VIF |
|---|---|---|---|---|---|
| constant | 0.063 | 0.750 | | 0.084 | 0.934 |
| GDP (LCU) | $1.01 \times 10^{-12}$ | 0.351 | 1.515 | 0.156 | 310.52 |
| GDP per capita | -0.01 | -0.372 | -2.104 | 0.057 | 181.65 |
| Remittan ce | $1.23 \times 10^{-10}$ | 0.30 | 1.458 | 0.170 | 244.81 |
| Tourist arrival | $1.86 \times 10^{-6}$ | 0.410 | 6.595 | 0.00 | 22.36 |
| CPI | 0.11 | 0.44 | 1.987 | 0.070 | 288.96 |
| Exchang e rate | -0.022 | -0.28 | -2.62 | 0.022 | 69.982 |
| Labour permit | $2.22 \times 10^{-6}$ | 0.30 | 3.40 | 0.005 | 45.030 |
| World GDP | $-2.28 \times 10^{-14}$ | -0.379 | -2.033 | 0.065 | 202.06 |
| PPP | 0.060 | 0.283 | 1.952 | 0.075 | 122.2 |

In the process of establishing a robust predictive model for air passenger demand, various combinations of independent variables were subjected to Multiple Linear Regression (MLR). The iteration yielded a statistically significant model, wherein the inclusion of the number of labor permits, tourist arrivals, and exchange rate as predictor variables demonstrated notable efficacy.

The resulting model exhibited a commendable coefficient of determination ( $R^2$ ) value of 0.986,shown in table5, signifying that approximately 98.6% of the variance in air passenger demand could be elucidated by the independent variables within the model.As shown in Table 6 The overall model demonstrated a favorable fit, as evidenced by a p-value below 0.05.Furthermore, the model's overall significance was underscored by an impressive F-statistic value of 411.14. This F-value, indicative of the model's overall significance, substantiates its robustness in explaining the variance in air passenger demand.

As shown in Table 7,upon individual examination of the predictor variables, the number of tourist arrivals emerged as the most influential, boasting a substantial

**Table 4:** Correlation Matrix

| | Pax movement | GDP (const LCU) | GDP per capita | Remittance | Tourist arrival | CPI | Exchange rate | Labour permit | World GDP | PPP |
|---|---|---|---|---|---|---|---|---|---|---|
| Pax movement | 1 | | | | | | | | | |
| GDP (const LCU) | 0.978 | 1 | | | | | | | | |
| GDP per capita | 0.992 | 0.985 | 1 | | | | | | | |
| Remittance | 0.990 | 0.998 | 0.991 | 1 | | | | | | |
| Tourist arrival | 0.926 | 0.902 | 0.924 | 0.889 | 1 | | | | | |
| CPI | 0.985 | 0.988 | 0.985 | 0.992 | 0.893 | 1 | | | | |
| Exchange rate | 0.893 | 0.913 | 0.896 | 0.916 | 0.795 | 0.945 | 1 | | | |
| Labour permit | 0.841 | 0.791 | 0.826 | 0.853 | 0.622 | 0.810 | 0.711 | 1 | | |
| World GDP | 0.956 | 0.948 | 0.954 | 0.957 | 0.851 | 0.928 | 0.780 | 0.893 | 1 | |
| PPP | 0.987 | 0.974 | 0.984 | 0.991 | 0.870 | 0.988 | 0.914 | 0.864 | 0.947 | 1 |

factor of 0.520 when all other variables were held constant. Furthermore, the statistical significance of all variables, including the number of tourist arrivals, exchange rate, and number of labor permits, was consistently recorded at levels below 0.05, underscoring their significant contributions to the model.The examination of multicollinearity concerns revealed VIF values below 5 for all variables, affirming the absence of issues pertaining to multicollinearity among the predictor variables.

In summary, the developed MLR model, anchored by the number of labor permits, tourist arrivals, and exchange rate, not only demonstrated high explanatory power but also showcased the individual significance of each variable, contributing valuable insights into the dynamics of air passenger demand. The equation for the model can be written as:

$$y = -1.25 + (2.414 \times 10^6)x_1 + (0.01)x_2 + (2.624 \times 10^6)x_3 \tag{3}$$

where,

$y$ = Air passenger demand (in millions)

$x_1$ = number of tourist arrivals

$x_2$ = Exchange rate

$x_3$ = number of labour permit

**Table 5:** Model Summary-Model 2

| R | $R^2$ | Adjusted $R^2$ | Std. error | D watson |
|---|---|---|---|---|
| 0.993 | 0.986 | 0.983 | 0.152 | 1.023 |

**Table 6:** ANOVA-Model 2

| model | S S | DF | Mean sqr | F | sig |
|---|---|---|---|---|---|
| Regression | 28.72 | 3 | 9.57 | 411.14 | 0.00 |
| Residual | 0.419 | 18 | 0.023 | | |
| Total | 29.14 | 2 | | | |

**Table 7:** Model Coefficients-Model 2

| Model | unstandard coef | standard coef | t | sig | VIF |
|---|---|---|---|---|---|
| constant | -1.248 | | -5.449 | 0.00 | |
| Tourist arrival | $2.41 \times 10^{-6}$ | 0.53 | 11.287 | 0.00 | 2.76 |
| Exchange rate | 0.017 | 0.22 | 4.211 | 0.001 | 3.42 |
| Labour permit | $2.62 \times 10^{-6}$ | 0.354 | 8.74 | 0.00 | 2.05 |

Figure 1 gives the idea about the normality of residuals where almost every residual closely follow a straight line from lower left to upper right corner which verifies the normality assumption of the model.

Figure 2 shows the points on the scatter plot between the dependent variable and independent variables forms an approximately straight line which verifies the linearity assumption for the model.

Figure 3 shows the residual plot where points on the graph does not form any pattern and are randomly distributed which implies the homoscedasticity assumption of the model is preserved.

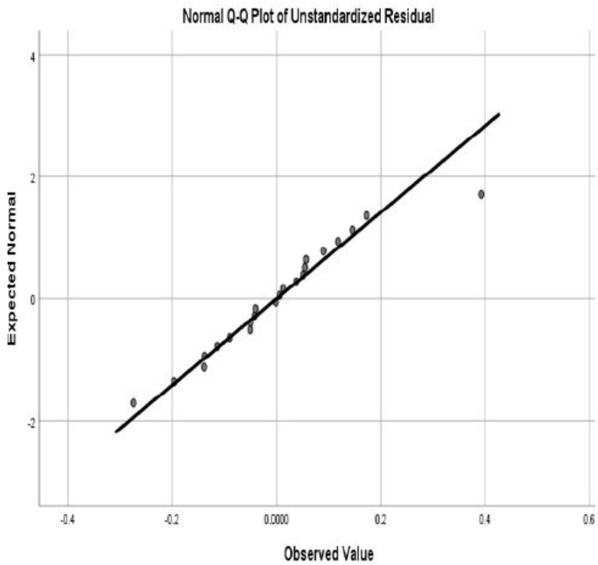


**Figure 1:** Normal Q-Q plot

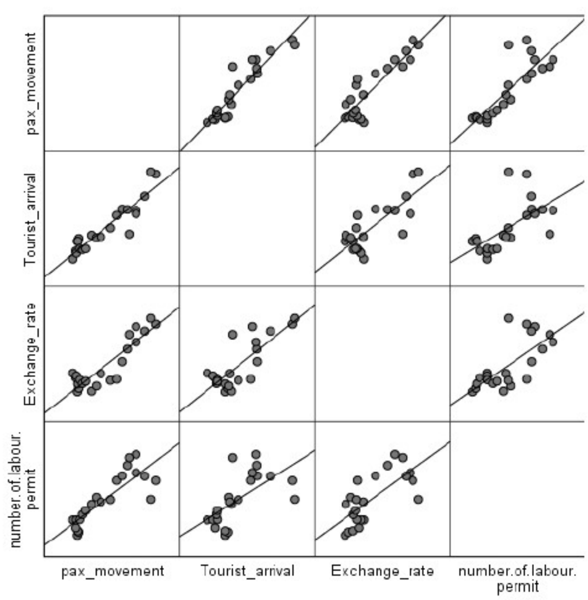


**Figure 2:** Scatterplot of dependent and independent variables

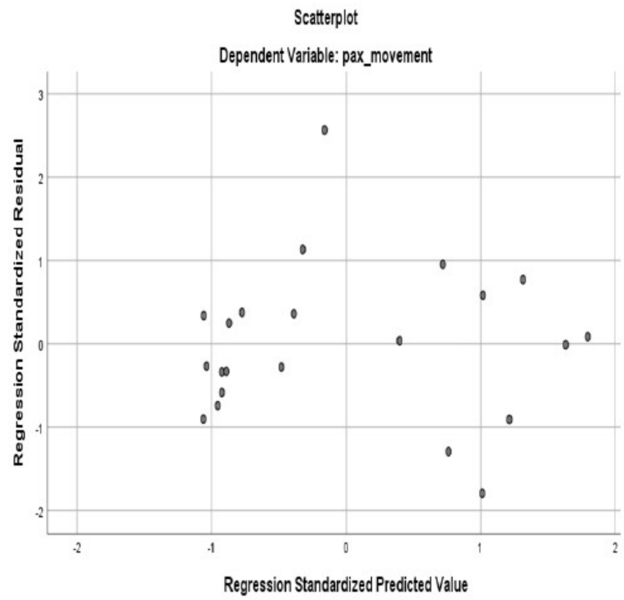


**Figure 3:** Scatterplot of residualsl

## 5. Conclusions and Recommendations

### 5.1 Conclusions

The study undertook a comprehensive analysis of the demand for international air passenger traffic at TIA, employing various socioeconomic variables. Initially, a range of factors, including GDP, tourist arrivals, CPI, and others, were considered in the modeling process. However, models incorporating variables like CPI, GDP, remittance, labour permits, tourist arrivals, and PPP were rejected due to observed high correlations between independent variables.

After rigorous exploration of different variable combinations, a robust demand model for air passengers emerged, featuring tourist arrivals, labor permits, and exchange rates as independent variables. This model demonstrated satisfactory results, providing valuable insights into the factors influencing international air passenger demand at TIA.

The variables 'number of labour permits' and 'number of tourist arrivals' represent dynamic figures influenced by a nation's socio-economic development. The increase in the number of labor permits may be attributed to social issues such as high unemployment rates and a lack of career development opportunities, leading individuals to seek opportunities abroad. Conversely, a rise in the number of tourist arrivals signifies the attractiveness of the country as a destination.Future studies could look more closely at how these factors interact, figuring out the detailed connections that might make our predictions about demand more accurate and useful. Understanding these dynamics is crucial for refining predictive models and gaining deeper insights into the socio-economic factors shaping international air travel demand.The findings contribute to a deeper understanding of the dynamics in this context and may inform policy and strategic decisions in the aviation sector.

### 5.2 Recommendations for further research

i) Only secondary data was used in the analysis. Primary data like income level of passenger, number of car/vehicle ownership etc , if incorporated in modelling might give better results.

ii) Impact of other airports like Gautam Buddha international airport, pokhara international airport could be considered for further studies to get accurate predictions.

## References


[1] M. Wang and H. Song, "Air travel demand studies: A review," *Journal of China Tourism Research*, vol. 6, no. 1, pp. 29–49, 2010.

[2] V. R. Rengaraju and V. T. Arasan, "Modeling for air travel demand," *JOURNAL OF TRANSPORTATION ENGINEERING*, vol. 118, no. 3, 1992.

[3] M. G. Karlaftis, K. G. Zografos, J. D. Papastavrou, and J. M. Charnes, "Methodological framework for air-travel demand forecasting," *JOURNAL OF TRANSPORTATION ENGINEERING*, vol. 122, no. 2, 1996.

[4] E. Erraitab, "An econometric analysis of air travel demand: The moroccan case," *European Scientific Journal*, vol. 12, no. 7, 2016.

[5] S. Y. Abed, A. O. Ba-Fail, and S. M. Jasimuddin, "An econometric analysis of international air travel demand in saudi arabia," *Journal of Air Transport Management*, vol. 7, pp. 143–148, 2001.

[6] D. P. Bastola, "Air passenger demand model (apdm): Econometric model for forecasting demand in passenger air transports in nepal," *International Journal of Advanced Research and Publications*, vol. 1, no. 4, 2017.

[7] J. Dargay and M. Hanly, "The determinants of the demand for international air travel to and from the uk," https://www.researchgate.net/publication/228602669, Month 2001, (Accessed on 2022/11/28).

[8] *CAAN Report 2017/2018/2019-20*, Civil Aviation Authority of Nepal (CAAN),Babarmahal, 2017/2018/2019.

[9] W. Bank, "Economic indicators," https://data.worldbank.org/indicator, june 2019, ().

[10] *Nepal Labour Migration Report*, Ministry of Labour Employment and Social Security, Singadarbar, 2018/2019/2020.

[11] *Nepal Tourism Statistics*, Ministry of Culture Tourism & Civil Aviation,Singhadarbar, 2017/2018/2019.

## Appendix
## Sample of Collected Data

| Year | Pax movement | GDP | GDP per capita | Remittance | Tourist arrival | CPI | Exchange rate | Labour permit | World GDP | PPP |
|---|---|---|---|---|---|---|---|---|---|---|
| 1998 | 1.04 | 9.26E+11 | 210.61 | 6.75E+07 | 464000 | 50.63 | 65.98 | 7745 | 3.15E+13 | 16.47 |
| 1999 | 1.1 | 9.67E+11 | 214.11 | 8.35E+07 | 492000 | 54.4 | 68.24 | 27796 | 3.27E+13 | 17.32 |
| 2000 | 1.06 | 1.03E+12 | 229.49 | 1.11E+08 | 464000 | 55.75 | 71.09 | 35543 | 3.38E+13 | 17.16 |
| 2001 | 0.98 | 1.08E+12 | 246.73 | 1.47E+08 | 361000 | 57.25 | 74.95 | 55025 | 3.36E+13 | 17.14 |
| 2002 | 0.85 | 1.08E+12 | 244.72 | 6.78E+08 | 275000 | 58.99 | 77.88 | 104736 | 3.49E+13 | 17.39 |
| 2003 | 1 | 1.12E+12 | 252.4 | 7.71E+08 | 338000 | 62.35 | 76.14 | 105043 | 3.91E+13 | 17.97 |
| 2004 | 1.14 | 1.17E+12 | 286.16 | 8.23E+08 | 385000 | 64.13 | 73.67 | 106660 | 4.41E+13 | 18 |
| 2005 | 1.25 | 1.21E+12 | 315.81 | 1.21E+09 | 375000 | 68.51 | 71.37 | 139718 | 4.77E+13 | 18.6 |
| 2006 | 1.38 | 1.25E+12 | 346.95 | 1.45E+09 | 384000 | 73.25 | 72.76 | 165252 | 5.18E+13 | 19.26 |
| 2007 | 1.63 | 1.30E+12 | 391.38 | 1.73E+09 | 527000 | 74.91 | 66.42 | 204533 | 5.84E+13 | 19.15 |
| 2008 | 1.83 | 1.38E+12 | 470.46 | 2.73E+09 | 500000 | 82.33 | 69.76 | 249051 | 6.41E+13 | 20.27 |