Estimation of the origin-destination matrix from national road traffic data in Central Java Province using the least squares method

Wahyuningsih Tri*1, Ary 1, Syafi’i 1

Abstract. The rapid development of Central Java Province is reflected in its robust economic, social, and cultural sectors, as well as in the high population and travel activities. Sustainable infrastructure development can accelerate equitable growth comprehensively [1-5]. These activities were further intensified by opening several cross-Java toll roads at the end of 2018. It is important to note that establishing training and land-use centers, such as government administration centers, settlements, schools, hospitals, entertainment facilities, and shopping centers, are responsible for these activities. These centers generate a large volume of movement that eventually impacts the province's transportation system [6].

One of the major causes of traffic congestion is the generation and attraction of traffic around the right and left access toll road exits in land-use areas [7]. This problem has worsened with the increasing traffic in Central Java Province, resulting from population growth, tourist destinations, and the proliferation of culinary establishments (food centers) [8]. Traffic management is one of the methods employed based on the analysis of existing traffic flow generation and attraction movements [9-10].

In Central Java Province, the Trans Java toll road provided users additional route options previously limited to national and regional roads. This new toll road offers a time-saving and cost-effective alternative as it reduces fuel consumption [11].

Toll roads can potentially divert traffic flow from national or old/existing regional routes, thereby alleviating congestion [12]. The construction of toll roads undoubtedly affects the performance of the national road network. According to Samosir and Mulyono [13], the traffic performance of road networks changes with loading due to network changes or the addition of new road networks [14-15]. Therefore, as a policy maker, the government needs to carefully address transportation problems based on the magnitude of the anticipated movement.

A rise in movement can be an early warning sign of impending transportation problems, which must be addressed promptly [16]. Transportation planning is crucial as it helps in considering the movement requirement. One approach to estimating movement patterns between two destinations is the Gravity Model, which employs the Exponential-Negative resistance function and the Least Squares method for settlement.

Therefore, this study aims to determine future traffic volume by forecasting Origin-Destination (O-D) Matrix with a gravity model equation [17]. The government needs to analyze the effect of toll road networks on the performance of the national road network in both current and future years [12]. In addition, one of the current government's main concerns is road performance, which is influenced by governance and is a critical issue [18]. This research is important to determine future traffic volume using a forecasting Origin-Destination (O-D) Matrix with a gravity model equation [17].

1 Introduction

The rapid development of Central Java Province is reflected in its robust economic, social, and cultural sectors, as well as in the high population and travel activities. Sustainable infrastructure development can accelerate equitable growth comprehensively [1-5].

These activities were further intensified by opening several cross-Java toll roads at the end of 2018. It is important to note that establishing training and land-use centers, such as government administration centers, settlements, schools, hospitals, entertainment facilities, and shopping centers, are responsible for these activities. These centers generate a large volume of movement that eventually impacts the province's transportation system [6].

One of the major causes of traffic congestion is the generation and attraction of traffic around the right and left access toll road exits in land-use areas [7]. This problem has worsened with the increasing traffic in Central Java Province, resulting from population growth, tourist destinations, and the proliferation of culinary establishments (food centers) [8]. Traffic management is one of the methods employed based on the analysis of existing traffic flow generation and attraction movements [9-10].

In Central Java Province, the Trans Java toll road provided users additional route options previously limited to national and regional roads. This new toll road offers a time-saving and cost-effective alternative as it reduces fuel consumption [11].

Toll roads can potentially divert traffic flow from national or old/existing regional routes, thereby alleviating congestion [12]. The construction of toll roads undoubtedly affects the performance of the national road network. According to Samosir and Mulyono [13], the traffic performance of road networks changes with loading due to network changes or the addition of new road networks [14-15]. Therefore, as a policy maker, the government needs to carefully address transportation problems based on the magnitude of the anticipated movement.

A rise in movement can be an early warning sign of impending transportation problems, which must be addressed promptly [16]. Transportation planning is crucial as it helps in considering the movement requirement. One approach to estimating movement patterns between two destinations is the Gravity Model, which employs the Exponential-Negative resistance function and the Least Squares method for settlement.

Therefore, this study aims to determine future traffic volume by forecasting Origin-Destination (O-D) Matrix with a gravity model equation [17]. The government needs to analyze the effect of toll road networks on the performance of the national road network in both current and future years [12]. In addition, one of the current government's main concerns is road performance, which is influenced by governance and is a critical issue [18]. This research is important to determine future traffic volume using a forecasting Origin-Destination (O-D) Matrix with a gravity model equation [17].
2 Basic Theory

2.1 Gravity models

Transportation modeling is employed in planning to allocate the destination matrix on the road network, thereby generating traffic flow [19]. In transportation system planning and modeling, the origin-destination (O-D) matrix is a primary data component. The Origin-Destination Matrix is the main component of the data needed in planning and modeling the transportation system [20].

Several methods can be used to model the origin-destination matrix, including the synthetic aspect involving the gravity model [21-22]. One of the techniques used to develop the O-D matrix is the artificial method with the Gravity Model. The gravity model produces the law of gravity analogy, influenced by the masses of two objects and the distance [23]. In the context of transportation planning, gravity suggests that trip generation and accessibility from the origin to the destination zone determine movement. The trip generation is analogous to the mass of two objects, while accessibility is equated with their distance. The generation and attraction limits are balanced with the following Equation 1.

\[
T_{id} = O_i \cdot D_d \cdot A_i \cdot B_d \cdot f(C_{id})
\]

given:
\[
T_{id} = \text{Total movement from the origin to the destination zone}
A_i, B_d = \text{Balancing factor for each zone of origin to the destination}
O_i = \text{Total moves generated from the origin zone}
D_d = \text{Total moves attracted toward the destination zone}
\]

2.2 Least squares method

The unknown value of the \( \beta \) parameter in the Gravity model equation can be obtained using the Least Squares method. In this technique, the calculation minimizes the difference between the modeling results and the observational data. The Least Squares estimation method can be expressed as follows Equation 2 [2].

\[
\frac{\partial S}{\partial \beta} = f = \sum_{i=1}^{N} \sum_{d=1}^{N} \frac{2}{N} \left( T_{id} - \hat{T}_{id} \right) \frac{\partial T_{id}}{\partial \beta}
\]

The Least Squares have become a widely used technique for estimating loading [14]. This method assumes the absence of measurement errors in the calculations, but in reality, errors can occur and affect the results. To obtain a more accurate and reliable estimate, the Least Squares account for measurement errors [14].

2.3 Newton-Raphson calibration

Parameter \( \beta \), not obtained in Equation 2, can be realized using the Newton-Raphson Calibration method. This technique is completed by eliminating the Gauss-Jordan matrix. The equation used is as follows Equation 3.

\[
h = -\frac{\partial f}{\partial \beta}
\]

The value of \( h \) is used to get the following deals Equation 4.

\[
\beta_1 = \beta_0 + h
\]

2.4 Traffic forecasting

As part of the road facility design process, investment feasibility studies and the development of work documentation are essential. Meanwhile, determining a comprehensive and valid car transportation and distribution plan for a complex road network requires mathematical models and appropriate software applications [22]. To forecast traffic, EMME/3 software, the Gravity Model, and the growth factor were used to determine the expected level of attraction and generation in the coming year.

2.5 Growth factor

Several technical criteria are required in road planning to optimize planning results and anticipate social or technical issues, including considering traffic growth. Traffic analysis can provide sufficiently accurate planning data in developing countries like Indonesia [24].

2.6 Statistical test indicator

The coefficient of determination is used to determine the degree of similarity between traffic flow modeling results and traffic count. The value of \( R^2 \) is derived from linear regression analysis (Equation 5), and the validity level of the model is determined as presented in Table 1.

\[
R^2 = \frac{\sum_i (Y_i - \bar{Y}_i)^2}{\sum_i (Y_i - \bar{Y})^2}
\]
### Table 1. Coefficient Range Determination (R²)

<table>
<thead>
<tr>
<th>Mark R²</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.80–1.00</td>
<td>very high</td>
</tr>
<tr>
<td>0.60–0.80</td>
<td>high</td>
</tr>
<tr>
<td>0.40–0.60</td>
<td>high enough</td>
</tr>
<tr>
<td>0.20–0.40</td>
<td>low</td>
</tr>
<tr>
<td>0.20–0.40</td>
<td>Very low</td>
</tr>
</tbody>
</table>

### 3 Study methods

#### 3.1 Study location

This study focuses on toll and national roads that traverse Central Java Province. The zoning system is based on districts/cities within Central Java Province, totaling 35 internal zones, each representing a district/city. Furthermore, nine external zones are considered, which include areas outside the regencies/towns but have access to the national and toll roads in the region. These external zones encompass several West Java, East Java, and D.I. Yogyakarta locations (Fig. 1).

#### 3.2 Data analysis method

The research was conducted using quantitative methods by analyzing all traffic movements on national roads and toll roads in Central Java Province. Federal roads in Central Java Province based on the Decree of the Minister of Public Works and Public Housing Number 290/KPTS/M/2015 totaling 237 roads along 1,518.05 Km and toll roads in Central Java Province as many as 21 roads along 306.05 Km. The Gravity model can estimate the distribution of future movements. In the equation of the Gravity model, there are parameter values to estimate the distribution of activities in the coming year. Gravity models have been widely used for commuter travel and can describe behavior that follows a pattern similar to Newton's famous law of gravity [21], [25]. The instrument as a tool to obtain data is carried out by entering traffic count data, side obstacle data, road network data, and calculation of road capacity and travel time following the standards of the Indonesian Road Capacity Manual [26]. Then make a road network database and estimate the Origin-Destination (O-D) matrix with EMME/3 software with a gravity model. Case studies were conducted to evaluate the magnitude of the movement pattern of the origin of the destination with the Gravity Model approach, where the obstacle function used is Exponential-Negative with completion using the Least Squares method.

#### 3.2.1 Database processing

The national road segment data were first processed to create a road network in the EMME/3 software. Then zones are grouped based on cities/regencies, and the center of each is represented by its centroid. Furthermore, the data on the coordinate values of the road sections according to ArcGIS and the calculation data for road capacity, speed, travel time, and traffic volume are inputted through the network editor. Preliminary matrix data are also required to estimate the matrix of traffic (O-D). This matrix assumes that each cell contains a value of one for movement between the same zone of origin and destination and zero for all other cells.

![Fig. 1. The study area in Central Java Province.](image-url)
3.2.2 Origin-destination (O-D) estimation results by EMME/3

After obtaining the database, it was inputted into the EMME/3 software, and the running process commenced. This process generates the MAT results of EMME/3 for 2022. The validity of the O-D matrix was tested to determine the accuracy level.

3.2.3 Cost matrix (Cid) in 2022

Aside from the movement matrix, another result obtained from the running process is the Cost matrix (Cid). The value of the cost matrix was obtained through the same steps employed in obtaining the movement matrix, which requires a single iteration process to gather the data.

3.2.4 Beta parameter calibration (β)

The calibration of the β parameter in the Exponential-Negative resistance function is done using the Newton-Raphson calibration method. This method involves repeating the calculation until a fixed β value is reached. The repetition process is performed using the Matlab program.

3.2.5 Origin-destination (O-D) estimation of modeling results in 2022

The O-D matrix estimation from the generation-attraction limit Gravity model is obtained by inserting the β value from the calibration result into Equation [1]. This process produces the 2022 O-D matrix Gravity Model.

3.2.6 Imposition of traffic flow to the road network

The traffic flow on the road is determined by loading O-D Matrix onto the road network using the EMME/3 software and applying the Wardrop Equilibrium principle. The result of this loading is the flow of traffic on the road.

3.2.7 Validation test

In the validity test of modeled traffic flow, the results are compared with those observed in the field using a statistical test known as the Coefficient of Determination (R²). The difference in the magnitude between the two traffic flow results is then assessed to determine its significance.

3.2.8 Origin-destination (O-D) of 2040 modeling results

The O-D Matrix estimation results for the 2040 Gravity model can be obtained through a forecasting process. This process involves using a growth factor (i%) value as a multiplier for the estimated Oi and Dd values. After getting the Oi and Dd values, the Cid value can be determined by repeating the same process employed for the current year. The Oi, Dd, and Cid values are then applied as a prior matrix, and the same step can be repeated to obtain the 2022 O-D Gravity Model. The O-D matrix in the 2040 Gravity Model results is then assigned to the road network sections. From the results of this assignment, it is possible to determine whether the road segment requires further evaluation and handling.

4 Results and discussion

4.1 Total movement of the road network

The 2022 Central Java Province Destination Origin Matrix is based on calculating traffic flow data (Matrix Estimation by Traffic Count) using EMME software, obtained by inputting the prior matrix through the prompt console. After inputting the previous matrix data, the program run process is carried out. In the EMME/3 program, the author stores the matrix of the results of the last run program process in the mf3 matrix table (full matrix 3). Once saved, the mf3 is moved into Microsoft Excel software to make it easier to analyze movements in all zones. The total movement of the road network in 2040 can be seen in Fig. 2.

![Fig. 2. Total movement of the road network in 2040](image)

The following are the data analysis results from the O-D matrix for Central Java Province in 2022 and 2040, indicating the number of movements on the road network, including toll routes.

<table>
<thead>
<tr>
<th>Year</th>
<th>Movement Total (PCU/hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2022</td>
<td>50,295</td>
</tr>
<tr>
<td>2025</td>
<td>59,629</td>
</tr>
<tr>
<td>2030</td>
<td>79,191</td>
</tr>
<tr>
<td>2035</td>
<td>105,171</td>
</tr>
<tr>
<td>2040</td>
<td>139,674</td>
</tr>
</tbody>
</table>

Table 2. The total movement every five years until 2040
4.2 Movement patterns zones

The division of an area is based on administrative boundaries in the form of cities/districts. The division of zones in Central Java Province is 44, with 35 internal zones (all cities/ regencies in Central Java Province) and nine external zones. Each zone has one representative zone center called a centroid, which is then connected to one of the road network nodes (nodes) with a link (centroid connector).

After conducting a calculation analysis on the total movement between zones in 2022, it was found that the internal-internal zone experienced the largest sign, totaling 37,528 PCU/hour. The highest generation in this area occurred in zone 702, which is Kendal City. The external zone with the most significant generated-attraction movement value was in the Kulonprogo Regency between Central Java Province and D.I Yogyakarta. The total movement of the road network in 2040 can be seen in Fig. 3.

![Fig. 3. The total movement zones in 2022 and 2040.](image)

The results of the new O-D matrix from the estimation process in 2022 can be seen in the Fig. 4.

![Fig. 4. The total movement zones in 2022 and 2040.](image)

In 2040, the total movement amounts to 139,674 PCU/hour, and the biggest movement generation is from internal training, which is 104,218 PCU/hour. The highest generation value in 2040 still originates from Kendal City, while the external signal recorded the lowest figure at 2,502 PCU/hour. The table below illustrates the results of the new O-D matrix obtained from the estimation process in 2040.
5.2 Recommendation

Based on the analysis and discussion that has been described previously, the following recommendations were proposed:

- Toll roads can be constructed in areas outside the center of the capital city or regency to divide internal and external movement.
- The movement generation on land use can be further analyzed and divided into internal and external movements.
- Further study is needed to examine the movement generation and attraction of movement.
- Internal movements within the region will increase by 66,690 PCU/hour in 2022 and 2040.
- External movement is projected to increase by 89,379 PCU/hour between 2022 and 2040.
- Consequently, Kendal is expected to experience the expected to increase by 10,317 PCU/hour, 10,770 PCU/hour, 9,718 PCU/hour, 10,039 PCU/hour, 10,317 PCU/hour, 10,039 PCU/hour, 9,718 PCU/hour, 10,770 PCU/hour, 10,317 PCU/hour, 10,039 PCU/hour, 9,718 PCU/hour.

References