Manila Journal of Science ***16(2)*** *2023, pp. 1–32*

Designing a Regional Railway Network Using an Improved Gravity Model and Graph Theory Approach

Joana Feliz M. Gayeta,
John Patrick D. Jamisola, and Yvette Fajardo-Lim*
Department of Mathematics and Statistics, De La Salle University, Manila, Philippines

* Corresponding Author: yvette.lim@dlsu.edu.ph

ABSTRACT

Railways have been one of the most efficient modes of transportation. They enable economic progress by mobilizing people and goods. With this, the researchers designed a regional railway network using graph theory with weighted nodes. The weight of each node will be obtained using the grey correlation coefficient of the following variables: total population, distance, perceived passengers, annual income, and the number of tourists. These factors will give an improved gravity model that will be used as the road weight between two cities. The layout for the regional railway network will be obtained using Kruskal's algorithm. The model will be applied to Region III (Central Luzon) of the Philippines to verify its validity.

Keywords: railway, network, graph theory, Kruskal's algorithm, gravity model, grey correlation

INTRODUCTION

Public transportation has always been one of the major contributors to a country's economic growth. It enables economic progress by mobilizing people and goods and maximizes the accessibility of a community to other areas. There are many kinds of public transportation that a country can have (e.g., buses, streetcars, trolleys, etc.), but the most ideal out of all of these is having a railway transit (Severino et al., 2021).

Although it is quite costly to build and maintain one, the advantages and benefits of having a railway system are almost never-ending. Compared to other modes of transportation, railways are faster and more dependable since they are not easily affected by changing weather and traffic jams. Their fixed routes and jam-packed schedule make it easier for people to travel from one place to another in a short period of time. Their carrying capacity is also quite large, which is beneficial in transporting bulk quantities of goods. In America alone, the impact of railway transportation is very evident since it helped consumers save billions of dollars every year and lessen the pollution and energy consumption, reduces exorbitant costs of highway development and maintenance for taxpayers, and provides more jobs to the people (Association of American Railroads, 2021). Looking at these, it can be easily said that building a convenient railway transport system is an important step in promoting coordinated economic development between cities.

Speaking of economic development, among the ASEAN countries, the Philippines is known for having one of the highest economic growth rates in the region

(Mercado, 2020). But as the population in the country increases over time, the traffic congestion and mobility of the people are affected, which causes a huge economic loss. According to the research conducted by the Japan International Cooperation Agency (JICA) in 2018, the country is losing an approximate ₱3.5 billion a day and is projected to lose ₱5.4 billion a day in 2035, due to traffic congestion in Metro Manila (CNN Philippines, 2018). On top of that, carbon dioxide (CO_2) emissions from the traffic affect the air condition, which causes health problems for the people living in the city. Now, these problems are not that new for the Philippines, especially in Metro Manila. Ever since the 1970s, cities in Metro Manila have been experiencing traffic jams causing local government units to implement traffic rules such as number coding and truck bans to lessen the congestion, but it was still not effective. Thus, the renovation and the construction of new railways were pushed through. After the construction of LRT1 in 1984, the congestion was lessened and all seemed to be going smoothly, but as the 1990s entered, the said railway reached its capacity due to a large number of passengers. Thus, everything went back to square one. Renovations, extensions, and the creation of new railway lines are very much needed (Orbon & Dungca, 2015). The same thing also happened to the Philippine National Railway (PNR). After undergoing some major renovations ever since it started its operation in the 1890s, the PNR brought convenience to the Filipino people for a few years, but after some time, its efficiency and railway coverage have gradually declined because of the continued neglect from the LGUs and damage from natural calamities (Philippine National Railways, n.d.). This cycle keeps going on and on until now, mainly because most of the railways that were created are only centered in Metro Manila (Asian Development Bank, 2012). To solve this issue, an infrastructure development project of the government called the “Build, Build, Build” policy was launched last 2017 under the Duterte administration. This involves the plan of creating several road networks, airports, seaports, bridges, and of course railway networks. Some of the proposed railway projects include the Malolos–Clark railway project that connects Malolos to Clark economic zone and Clark International Airport (CIA) in Central Luzon, Philippines, and the North–South Railway Project that connects the National Capital Region (NCR) and Legaspi City, the capital of Albay Province (Philippine News Agency, 2021).

Now, constructing railways is not just costly; it also needs some time and much effort since a lot of planning and research is needed to ensure that the railways will function at their best performance (MacKechnie, 2018). Since building railways is not common, a lot of methods and research are already conducted on how to efficiently create railway networks, but the problem is that these methods are customized to fit the countries of the respected researchers. Some of these include the study conducted by Zhang, Yong-sheng, Zeng, and Yin (2019), which is centered on the rail rapid transit trunk line and its effect on the compactness of urban agglomeration in China. They used some concepts of graph theory and two models, which are the rail rapid transit trunk line model and the urban agglomeration traffic compactness model. These models were applied in the West Triangle Economic Zone in order to verify their effectiveness. There are also studies that involve the use of a hybrid teaching-learning-based optimization algorithm to utilize and optimize an urban railway line in China (Liu et al., 2021) and the creation of a mathematical model and a genetic algorithm to solve the train timetabling problem in Germany (Arenas et al., 2015).

Other than designing the layout of railway networks, it is also important to check the vulnerability of an existing or a generated design layout of a railway network. Guze (2019) discussed in her study the different bottlenecks of a transportation network. Here she used the concepts of dominating sets, to determine which nodes and edges are prone to traffic flow problems. These concepts would be essential in designing a transportation network since these would determine all the types of disturbances that can occur during an operation. Furthermore, addressing the long-term assessment of constructing a railway network should also be considered. Canca, Andrade-Pineda, De-Los-Santos, and González (2021) created a model that can be used in the construction problem of an urban railway network. This model can be used to gain insight into the long-term cost of the construction project and the revenue. This is important to assess if the revenue will be outperformed by the operation cost in the long term.

In the Philippines, there is a newly established research and training center that focuses on the development of management, operation, and maintenance of railways. This center is called the Philippine Railway Institute (PRI) and was established with the help of the Japanese government last 2019 (Pateña, 2019). But despite having this research institute, railway planning and development are still not done by Filipino professionals; thus, help from outsiders, such as the Japanese companies, is still needed (Sumito Corporation, 2020).

With all of these, the researchers identified three major problems that this paper hopes to address. First of which is the traffic congestion in Metro Manila that causes huge economic loss for the country (CNN Philippines, 2018). By creating a regional railway network, citizens (of different regions) would realize that there is no need to stay in Metro Manila since they too will have their own accessible public transportation in their respective provinces which could lead the region to have more economic opportunities in the future. Aside from this, this research hopes to help PRI in establishing methods for creating railway transit plans since most of the research articles and methods that the railway planners are using were conducted and developed outside the Philippines. Although there is nothing wrong with it, the methods developed by other professionals are somewhat “custom-fit” to their countries, which means that there is a high percent chance that it will not be as effective in the Philippines. To fully maximize the potential of railway networks in the Philippines, it is only fitting to conduct a research based on the data gathered in the country by using a spatial interaction model called the gravity model and applying the concepts of graph theory to design a process of creating a regional railway network suited for the country.

This paper aims to show that an improved gravity model that will consider perceived passengers, annual income, and the number of tourists in addition to total population and distance will prove to be useful to assign weights to each city represented as nodes of a graph. These weighted nodes will be used to design the railway network using Kruskal’s algorithm. The model will be applied to Region III (Central Luzon) of the Philippines to verify its validity.

The results of this study would help railway planners save time in planning the overall possible regional railway network since using the improved gravity model would help them easily determine how strong the interaction between two cities is, thus knowing which cities the network should pass through. With this, they could easily see if their plans are feasible and take out unnecessary process(es) that would take too much of their time. Furthermore, conducting this research would also be beneficial for future plans and projects of the local governments in certain provinces or cities in the Philippines. They could recommend the process used in this research to their own railway planners if they are planning to build a railway network within their own province since the data and variables that will be used in conducting this research are readily available from the government sites of the country. Finally, this research can be the key to opening huge opportunities for the different regions in the country. This is because once the region has a means of constructing its own railway network, it could easily implement it and there is a high percent chance that a lot of companies and organizations would be interested in building infrastructures within the region.

METHODOLOGY

The study will be based on a gravity model that considers the population size of two places and the distance separating them (Rosenberg, 2019). These two variables will be used to predict or estimate the volume of flow or spatial interaction between or among places. The expectation is that the weight of or spatial interaction has a direct relationship with the population total of the locations and an inverse relationship with the distance of the said locations (Overman, 2009). Thus, if the population totals of two cities are large, the weight of spatial interaction between the cities will increase, but if the distance of the cities is larger, the level of interaction is expected to decrease. The gravity model is given below (Wheeler, 2005).

$$W_{ij} = k\left(\frac{P_i P_j}{D_{ij}^r}\right) \quad (1)$$

where

W_{ij} is the estimate of the spatial interaction between place of origin i and place of destination j;
k is a constant;
P_i, P_j are the population sizes of the places of origin and destination, respectively.
D_{ij} is the distance between i and j; and
r is the exponent of distance.

The authors will use an improved one, which will include the gross domestic product (GDP) of a city. However, in the Philippines, GDP is computed at a national and regional level only. Data for domestic products per city are not available. Furthermore, the interaction between two cities is not only limited to their population and GDP, but this also includes tourism, the flow of the population from one city to another, etc. (Zhang et al., 2019). Therefore, the researchers selected four factors that can be quantitatively analyzed, including the resident population, the annual income, the perceived passengers, and the number of tourist visits. The resident population and the perceived passengers of the cities are considered since these affect the number of commuters/passengers that would use the proposed network layout. In addition to this, instead of using the GDP, the annual income of the cities was also considered since it describes the financial capability of a city to provide developmental programs (Exec. Order No. 249, 1987). Lastly, the number of tourists was also deemed to affect the weight of interaction of the cities since it increases the mobility and the flow of people going inside or outside of the cities.

To improve the gravity model for the purpose of this study, the following variables will be needed: total population, perceived passengers, and annual income. These variables were collected through the website of the Philippines Statistics Authority (PSA), psa.gov.ph. This website provides the data needed for each variable at the region and city levels. The number of tourists was also collected from the website of the Department of Tourism (DOT). Both city level and regional level were collected. The distance between the two cities was also needed in the improved gravity model. This data was collected with the help of Google Maps. Table 1 shows the summary for retrieving the data.

Table 1. Summary for Retrieving the Data

Variable	Retrieved From	Remarks
Population	2020 Regional Social and Economic Trends from PSA	Collected per city level and regional level.
Perceived passengers	Census of Population from PSA	Only retrieved age 15–64, which is considered as the working population. Data were collected at city and regional levels.
Distance	Google Maps	Only cities that are connected by a national road were calculated.
Annual income	2020 Regional Social and Economic Trends from PSA	Collected per city level and regional level.
Number of tourists	2015 Tourist Demand Statistics—Regional Travelers from DOT	The number of tourists in the province was used as the basis for the number of tourists in each city located in that province.

Note. DOT = Department of Tourism; PSA = Philippine Statistics Authority.

In this paper, the researchers have determined these five possible factors that could affect the outcome of the regional railway network layout. The grey correlation analysis would be applied to find the index weight of each factor that will be used to complete the formula for the improved gravity model. The calculation steps to determine the weight coefficient are as follows (Liang et al., 2020):

1. Use the initial value to unify the dimensions of each factor sequence to minimize incommensurability caused by disparate dimensions:

$$Y'_{ij} = \frac{Y_{ij}}{Y_{0j}} \ , \ i = 1,2,\dots,n \ ; \ j = 1,2,\dots m \qquad (2)$$

where

Y'_{ij} is the resulting analysis sequence;
Y_{ij} is the initial value of the analysis sequence;
Y_{0j} is the reference sequence, usually a real number greater than 0;
m is the number of variables/factors (in this case $m = 4$); and
n is the number of observed years in the data set.

2. Compute the relationship coefficient between the analysis sequence and the reference sequence:

$$t_{ij} = \frac{min_n min_m \left|Y'_{0j} - Y'_{ij}\right| + \rho max_n max_m \left|Y'_{0j} - Y'_{ij}\right|}{\left|Y'_{0j} - Y'_{ij}\right| + \rho max_n max_m \left|Y'_{0j} - Y'_{ij}\right|} \qquad (3)$$

where

t_{ij} is the grey correlation degree of analysis sequence $Y'_{ij} (i = 1,2,\dots,n; j = 1,2,\dots m)$ with respect to reference sequence $Y'_{0j} (j = 1,2,\dots m)$; and
$\rho \in (0,1)$ is the resolution coefficient.

3. The correlation degree of the index is determined by the correlation coefficient since the values of $t_{1j}, t_{2j}, \dots, t_{nj}$ signify the correlation between the reference value of i ($i = 0,1,2,\dots,n$) analysis sequence and the j^{th}

index value, that is, if it reflects the proportion of the j^{th} index in the whole index space:

$$p_j = \frac{1}{n}\sum_{i=1}^{n} t_{ij} \quad (4)$$

where p_j is the correlation degree of the index j.

APPLICATION, RESULTS, AND DISCUSSION

The researchers decided to utilize the available data of Region III Central Luzon to apply and verify the method presented. The researchers first created the initial railway network of the model by identifying which cities are connected to one another. To do this, the current road network of Region III Central Luzon from the website Accu-map (n.d.) was observed (see Figure 1). The red lines represent the major highways and expressways along the region while the yellow lines represent the provincial and city-level roads.

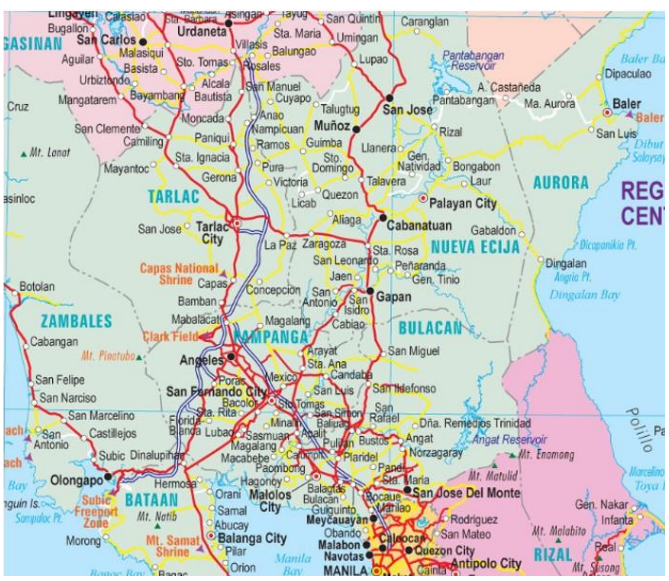


Figure 1. Road network of Region III from Accu-map.com.

With this current road network, the researchers looked for roads at or above the city level that connect two cities from the list. If there is an existing road, a virtual edge is connected from that city to another. In an event that there is more than one road at or above the city level that connects the two cities, the road having the shortest distance was selected. For example, there are three roads connecting Tarlac City and Muñoz, all of which are at or above city level, so we select the road that has the shortest distance, which is 48.7 km. Following these conditions, the initial railway network is shown in Figure 2, with 14 vertices and 31 edges, which is an undirected graph having E1 to E31 edges.

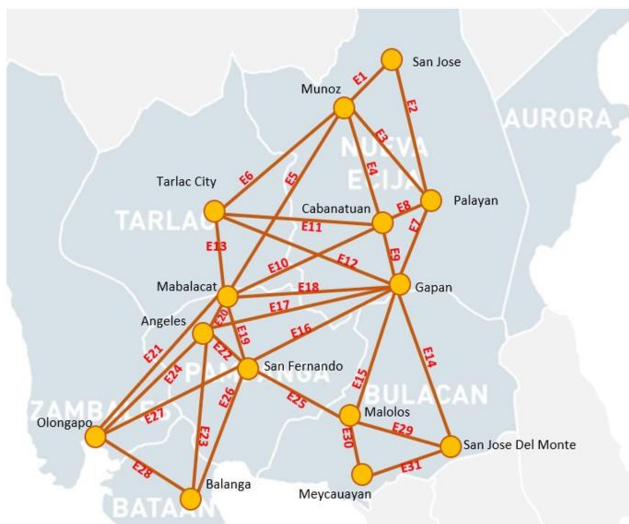


Figure 2. Initial railway network of Region III.

The table below shows the list of all the edges that were identified in Figure 2. The distance of each edge, in kilometers, was also collected since this would be used in the latter part of the paper. The distance indicated in the table is the shortest distance that one can travel from one city to another using a road at or above the city level.

Table 2. Distance Between Cities

Edge Number	Name of City 1	Name of City 2	Distance
E1	San Jose	Muñoz	20.9
E2	San Jose	Palayan	37.9
E3	Muñoz	Palayan	54.1
E4	Muñoz	Cabanatuan	37.1
E5	Muñoz	Mabalacat	85.3
E6	Muñoz	Tarlac	48.7
E7	Palayan	Gapan	50.4
E8	Palayan	Cabanatuan	19.8
E9	Cabanatuan	Gapan	31.6
E10	Cabanatuan	Mabalacat	73.5
E11	Cabanatuan	Tarlac	50.3
E12	Tarlac	Gapan	65.4
E13	Tarlac	Mabalacat	30.3
E14	Gapan	San Jose Del Monte	74.9
E15	Gapan	Malolos	64.8
E16	Gapan	San Fernando	62.4
E17	Gapan	Angeles	62.8
E18	Gapan	Mabalacat	88.6
E19	Mabalacat	San Fernando	25
E20	Mabalacat	Angeles	9
E21	Mabalacat	Olongapo	76.9
E22	Angeles	San Fernando	12.5
E23	Angeles	Balanga	72.8
E24	Angeles	Olongapo	73.9
E25	San Fernando	Malolos	33.1
E26	San Fernando	Balanga	56.1
E27	San Fernando	Olongapo	61
E28	Olongapo	Balanga	47.6
E29	Malolos	San Jose Del Monte	30.9
E30	Malolos	Meycauayan	22.6
E31	San Jose Del Monte	Meycauayan	19

In addition to these, the data on the influencing factors in Region III were also gathered. The data on population and annual income came from the 2020 Regional Social and Economic Trends of Region III Central Luzon, while the data on perceived passengers came from the census of the population from PSA and was the sum of the population with the age of 15–64 since it is considered as the working population. The number of tourists per city, on the other hand, came from the DOT's official website. The gathered regional data are as follows:

Table 3. Data From Central Luzon in 2015–2018

Year	Population	Perceived Passengers	Annual Income	Number of Tourists
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
2016	11,461,611	7,465,177	13,318,594,748	4,150,586
2017	11,705,046	7,623,730	15,530,180,025	3,903,037
2018	11,948,480	7,782,284	16,493,373,741	4,486,364

The researchers decided to focus on the years 2015 to 2018 since these are prepandemic years. Furthermore, the data of the region, excluding the population and the perceived passengers, were complete, unlike the succeeding years. For the population and perceived passengers, the researchers decided to use the 2015–2020 annual growth rate of Region III from the Philippine Statistics Authority, which is 2.17%, to compute the value for the years 2016 to 2018. This is because the available data on population and perceived passengers were collected every five years.

As for the city level, the researchers also collected the data from the same sites as the regional level. Unfortunately, the data on the number of tourists that the DOT has are only at the province level and not the city level. To address this, the researchers divided the data of a certain province by the number of cities it has in the list (e.g., since there are five cities in the list that belong to Nueva Ecija, we divide it into five). Note that this process was done under the assumption that the data from DOT were the number of tourists who came to each province and stayed in the cities, not the municipalities, of the region. Below is the table of the consolidated data per city.

Table 4. Data From Central Luzon in 2015

Name of City	Population	Perceived Passengers	Annual Income	Number of Tourists
Angeles	411,634	277,574	1,436,713,978	412,829
Balanga	96,061	63,979	521,987,321	244,303
Cabanatuan	302,231	199,374	1,056,972,930	5,197
Gapan	110,303	71,668	454,728,278	5,197
Mabalacat	250,799	166,090	999,537,542	500,990
Malolos	252,074	172,135	754,753,646	27,085
Meycauayan	209,083	134,816	907,431,879	27,085
Muñoz	81,483	49,939	389,238,966	5,197
Olongapo	233,040	147,005	1,255,249,945	164,042
Palayan	41,041	25,071	272,329,409	5,197
San Fernando	306,659	206,158	1,191,068,312	500,990
San Jose	139,738	90,281	554,412,749	5,197
San Jose Del Monte	574,089	371,847	1,078,619,971	27,085
Tarlac City	342,493	221,706	1,060,385,689	47,696

With all the accumulated data, the researchers then calculated the grey correlation of the influencing factors to compute the weight that would be used to improve the gravity model. To do this, the researchers first minimize incommensurability caused by disparate dimensions of the data in Table 3. The values for each factor were divided by their respective initial value (i.e., the value of each factor in the year 2015). The ratio for these values would be the analysis sequence. Now, take A = (1,1,1,1) as the reference sequence since it is a real number greater than 0. Note that the reference sequence is dimensionless and the initial value for each factor sequence. The relationship coefficient between the reference sequence and the analysis sequence can be computed by subtracting the reference sequence and the analysis sequence and by following Formula 2 with a resolution coefficient of 0.5. Using Formula 3, the grey correlation coefficient for each factor can be computed. To compute the index weight, the grey correlation coefficient of each factor was divided by the sum of all grey correlation coefficients. The result for this was summarized in Table 5 (see Appendix A for the full calculation of the grey correlation coefficient of each factor and its index weight).

Table 5. The Grey Correlation Coefficient and Index Weight

Year	Population	Perceived Passengers	Annual Income	Number of Tourists
2015	1	1	1	1
2016	0.886	0.886	0.677	0.523
2017	0.796	0.796	0.394	0.665
2018	0.722	0.722	0.333	0.406
Grey correlation coefficient	0.851	0.851	0.601	0.649
Index weight	0.288	0.288	0.204	0.220

It can be observed that the values for the population and the perceived passengers were the same since their data on each factor in Table 3 were relatively close to each other, and thus, when the researchers made these factors dimensionless (i.e., by getting their ratio), the values computed were almost the same up to the thousandths decimal place (refer to Appendix A to see the full calculation of the grey correlation). Furthermore, the population and perceived passengers both had larger weights compared to the other factors. This indicates that these factors have a much more significant impact on the weight of interaction between two cities. Substituting the index weight calculated from the influencing factors (see Table 5) to Formula 4, the improved gravity model is now as follows:

$$W_{ij} = \frac{0.28834\left(\frac{P}{p_i+p_j}\right)+0.28834\left(\frac{G}{g_i+g_j}\right)+0.2035\left(\frac{E}{e_i+e_j}\right)+0.21972\left(\frac{T}{t_i+t_j}\right)}{\frac{1}{d_{ij}}} \quad (6)$$

where

i,j represents the two connected cities;

W_{ij} is the weight of interaction between city i and j;

P is the total population of the region;

p_i, p_j are the population of city i and j, respectively;
G is the total perceived passengers in the region;
g_i, g_j are the perceived passengers per city i and j, respectively;
E is the annual income of the region;
e_i, e_j are the annual income of city i and j, respectively;
T is the total number of tourists in the region;
t_i, t_j are the number of tourists of city i and j, respectively; and
d_{ij} is the distance, in kilometers, between city i and j.

With the now improved gravity model, the researchers then computed the weight of interaction of each edge by substituting the values of Table 2 and Table 4 into Formula 6 (see Appendix B for full computation). From here on, the edges will be denoted by $e_i, i = 1, \ldots, 31$ instead of E_i. The results of the computation are as follows:

Table 6. Weight of Edges

Edge Number	Name of City 1	Name of City 2	Weight of Edge	Rank
e_1	San Jose	Muñoz	2,263.71	24
e_2	San Jose	Palayan	4,365.47	29
e_3	Muñoz	Palayan	7,266.93	31
e_4	Muñoz	Cabanatuan	3,511.25	28
e_5	Muñoz	Mabalacat	1,950.05	22
e_6	Muñoz	Tarlac	1,561.33	19
e_7	Palayan	Gapan	6,180.20	30
e_8	Palayan	Cabanatuan	1,914.67	21
e_9	Cabanatuan	Gapan	2,947.44	27
e_{10}	Cabanatuan	Mabalacat	1,057.65	13
e_{11}	Cabanatuan	Tarlac	1,315.07	18
e_{12}	Tarlac	Gapan	2,022.19	23
e_{13}	Tarlac	Mabalacat	410.39	4
e_{14}	Gapan	San Jose Del Monte	2,665.83	25
e_{15}	Gapan	Malolos	2,859.24	26
e_{16}	Gapan	San Fernando	1,149.48	16
e_{17}	Gapan	Angeles	970.26	11
e_{18}	Gapan	Mabalacat	1,869.36	20
e_{19}	Mabalacat	San Fernando	334.90	3
e_{20}	Mabalacat	Angeles	103.74	1
e_{21}	Mabalacat	Olongapo	1,208.86	17
e_{22}	Angeles	San Fernando	133.47	2
e_{23}	Angeles	Balanga	1,093.37	14
e_{24}	Angeles	Olongapo	907.71	10
e_{25}	San Fernando	Malolos	468.19	6

e_{26}	San Fernando	Balanga	1,029.49	12
e_{27}	San Fernando	Olongapo	864.72	9
e_{28}	Olongapo	Balanga	1,102.39	15
e_{29}	Malolos	San Jose Del Monte	733.72	8
e_{30}	Malolos	Meycauayan	677.37	7
e_{31}	San Jose Del Monte	Meycauayan	458.64	5

Below are sample computations of the edges with the strongest and weakest weights of spatial interaction:

Edge 20 (e_{20}) connecting Mabalacat and Angeles:

$$W_{ij} = \frac{0.28834\left(\frac{P}{p_i + p_j}\right) + 0.28834\left(\frac{G}{g_i + g_j}\right) + 0.2035\left(\frac{E}{e_i + e_j}\right) + 0.21972\left(\frac{T}{t_i + t_j}\right)}{\frac{1}{d_{ij}}}$$

$$= \frac{0.28834\left(\frac{11,218,177}{250,799 + 411,634}\right) + 0.28834\left(\frac{7,306,623}{166,090 + 277,574}\right) + 0.2035\left(\frac{12,321,106,360}{999,537,542 + 1,436,713,978}\right) + 0.21972\left(\frac{3,596,097}{500,990 + 412,829}\right)}{\frac{1}{9}}$$

$$= \frac{0.28834(16.93481) + 0.28834(16.46882) + 0.20358(5.05740) + 0.21972(3.93524)}{0.11111}$$

$$= \frac{4.88310 + 4.74873 + 1.02961 + 0.86466}{0.11111} = 103.73594$$

Edge 3 (e_3) connecting Muñoz and Palayan:

$$W_{ij} = \frac{0.28834\left(\frac{P}{p_i + p_j}\right) + 0.28834\left(\frac{G}{g_i + g_j}\right) + 0.2035\left(\frac{E}{e_i + e_j}\right) + 0.21972\left(\frac{T}{t_i + t_j}\right)}{\frac{1}{d_{ij}}}$$

$$= \frac{0.28834\left(\frac{11,218,177}{81,483 + 41,041}\right) + 0.28834\left(\frac{7,306,623}{49,939 + 25,071}\right) + 0.2035\left(\frac{12,321,106,360}{389,238,966 + 272,329,409}\right) + 0.21972\left(\frac{3,596,097}{5,197.4 + 5,197.4}\right)}{\frac{1}{54.1}}$$

$$= \frac{134.29289}{0.01848} = 7,266.93145$$

Out of all the connected cities in the list, Mabalacat, Pampanga, and Angeles, Pampanga, have the strongest spatial interaction since they have the lowest weight. This would indicate that the resulting minimum spanning tree of the model would start on the edge connecting the nodes Angeles and Mabalacat followed by the succeeding edges. On the other hand, even though Muñoz and Palayan belong in the same province, they have the weakest spatial interaction since they have the highest weight. This would mean that the edge connecting the said cities has a high chance to not be included in the resulting minimum spanning tree since there are other edges that have a lower weight that would pass both Muñoz and Palayan.

The resulting weights will be ranked from lowest to highest to apply Kruskal's algorithm. Since e_{20} has the lowest weight, the minimum spanning tree will start from there followed by the second edge, which is e_{22}, and so on. Now if a loop is formed in the process, the edge causing the loop will not be included. Once all the nodes have been passed through, the minimum spanning tree is formed (see Appendix C for the full implementation of Kruskal's algorithm). The figure below shows the resulting minimum spanning tree:

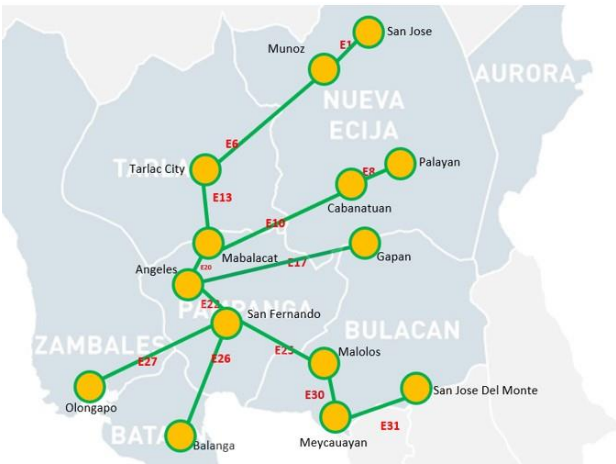


Figure 3. Resulting minimum spanning tree.

With the network layout above, the researchers came up with the final regional railway network for Region III Central Luzon. As a suggestion, the final network can be roughly divided into three lines with Angeles, Pampanga, as the center of the railway network as shown in the figure below.

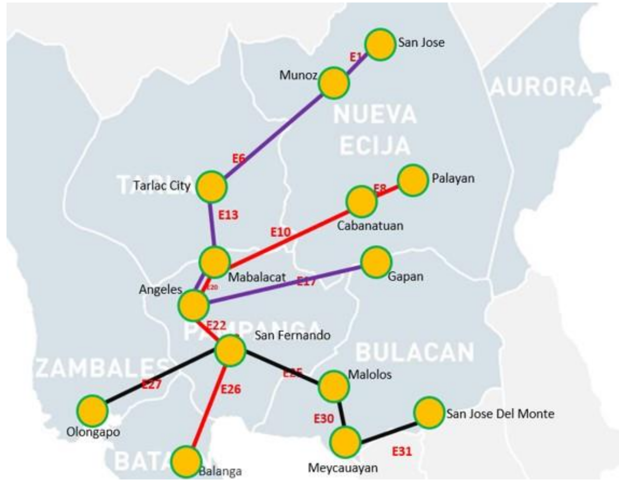


Figure 4. Final regional railway network of Central Luzon.

The three lines shown in Figure 4 were only a suggestion made by the researchers and do not necessarily mean that those are the only lines that can be created; the lines can still vary depending on the preferences of the railway engineers. Although dividing the network into different lines is not necessary, the researchers deemed it to be efficient because if it is not divided, the passengers need to follow a very strict time schedule every time they use the railway. For example, if the passenger from San Jose wanted to go to Balanga, the passenger cannot ride the train if it would go to Palayan, Gapan, Olongapo, or San Jose Del Monte and should wait for the train going to Balanga. But if there are different lines, such as the suggested lines, the passenger could just transfer lines. Furthermore, it can be said that this layout can be considered a reasonable route since it not only connects the cities based on its distance but also considered other factors such as the population, the perceived passengers, the annual income, and the number of tourists of each city.

CONCLUSION AND RECOMMENDATION

There is no one way in planning transit networks. With this, the researchers proposed a method that utilizes the mutual attraction between the cities in creating a regional transit network by considering four factors, namely, population, perceived passengers, annual income, and the number of tourists. The proposed method showed that adding factors to the gravity model that utilizes only distance and population would give a better layout for constructing a regional network.

The improved gravity model with the additional factors was formed by computing their grey correlation coefficient. This model was used to determine the strength of interaction between the cities after considering the four factors. The result after using the said model also serves as the weight of the edges for an undirected graph. With these weights, Kruskal's algorithm was used by the researchers to determine the resulting regional railway network. The model was applied to Region III of the Philippines to verify its validity.

After applying the model and seeing the result, the researchers recommend future researchers to add other relevant factors other than the four factors mentioned in this paper to help strengthen the effectiveness of the improved gravity model. Some factors that could be considered are the actual number of students and the working population since this paper only focused on the perceived passengers aging from 15 to 64. The number of projects on infrastructure and business opportunities per city would also be a great additional factor since it could affect the flow of people and goods in a certain city. Future researchers are also recommended to use accurate data for each factor to ensure the accuracy of the resulting minimum spanning tree. Furthermore, they could also add stations in municipalities of different provinces in the region since the research was only limited to the city level. Lastly, future researchers could apply the model to other regions to further check its accuracy and reliability.

REFERENCES

Accu-map. (n.d.). *Philippine Products & Services Page - Accu-map*. Retrieved June 8, 2022, from https://accu-map.com/philippine-products-services/philippine-products-services-product/

Arenas, D., Chevrier, R., Hanafi, S., & Rodriguez, J. (2015). *Solving the train timetabling problem, a mathematical model and a genetic algorithm solution approach.* 6th International Conference on Railway Operations Modelling and Analysis (RailTokyo2015), Mar 2015, Tokyo, Japan. https://hal.archives-ouvertes.fr/hal-01338609

Asian Development Bank. (2012). *Philippines: Transport sector assessment, strategy, and road map*. Retrieved June 5, 2022, from https://www.adb.org/sites/default/files/institutional-document/33700/files/philippines-transport-assessment.pdf

Association of American Railroads. (2021). *The economic impact of America's freight railroads*. https://www.aar.org/wp-content/uploads/2020/08/AAR-Economic-Impact-Fact-Sheet.pdf

Canca, D., Andrade-Pineda, J. L., De-Los-Santos, A., & González-R, P. L. (2021). A quantitative approach for the long-term assessment of Railway Rapid Transit network construction or expansion projects. *European Journal of Operational Research*, *294*(2), 604–621. https://doi.org/10.1016/j.ejor.2021.02.018

CNN Philippines. (2018). *PH loses ₱3.5B a day due to Metro Manila traffic – JICA*. Retrieved from https://cnnphilippines.com/transportation/2018/02/23/JICA-P3.5-billion-traffic.html

Department of Tourism. (n.d.). *Tourism demand statistics*. http://www.tourism.gov.ph/tourism_dem_sup_pub.aspx

Executive Order No. 249, s.1987. https://www.officialgazette.gov.ph/1987/07/25/executive-order-no-249-s-1987/

Guze, S. (2019). Graph theory approach to the vulnerability of transportation networks. *Algorithms*, *12*(12), 270. https://doi.org/10.3390/a12120270

Liang, T., Liu, H., & Tan, Y. (2020). Research on the gravity planning model of prefecture city rail transit network. *E3S Web of Conferences*, *145*, 02005. https://doi.org/10.1051/e3sconf/202014502005

Liu, F., Liu, C., Zhao, Q., & He, C. (2021). A hybrid teaching-learning-based optimization algorithm for the travel route optimization problem alongside the urban railway line. *Sustainability*, *13*(3), 1408. ResearchGate. https://www.researchgate.net/publication/348895543_A_Hybrid_Teaching-Learning-Based_Optimization_Algorithm_for_the_Travel_Route_Optimization_Problem_alongside_the_Urban_Railway_Line

MacKechnie, C. (2018). *What's the cost of a rail transit project?* LiveAbout. https://www.liveabout.com/rail-transit-projects-costs-2798796

Mercado, N. (2020). *PH remains one of fastest growing economies in Asia, says Salceda*. https://newsinfo.inquirer.net/1218842/ph-remains-one-of-fastest-growing-economies-in-asia-says-salceda

Orbon, G., & Dungca, J. (2015). The Philippine railway system. ResearchGate. https://www.researchgate.net/publication/319263143_The_Philippine_Railway_System

Overman, H. K. (2009). Regional science. In *International Encyclopedia of Human Geography* (pp. 279–285). https://doi.org/10.1016/B978-008044910-4.00737-9

Pateña, A. J. (2019). *DOTr, JICA approve curriculum for PH Railway Institute*. https://www.pna.gov.ph/articles/1063296

Philippine National Railways. (n.d.). *PNR in Philippine History*. Retrieved June 6, 2022, from https://pnr.gov.ph/about-contact-us/who-we-are/pnr-in-philippine-history/pnr-in-philippine-history.html

Philippine News Agency. (2021). *'Build, Build, Build' delivers, sets up infra for future*. https://www.pna.gov.ph/articles/1144279

Philippine Statistics Authority. (2021). *2020 regional social and economic trends*. http://rsso03.psa.gov.ph/sites/default/files/2020%20RSET.pdf?fbclid=IwAR2aWPD9hDD4avdQZJxpCOizL_ZSWcK9UMb0K6Eb3Mi7mJqaS-wZG2r82zY

Rosenberg, M., (2019). *Understanding the gravity model*. https://www.thoughtco.com/what-is-the-gravity-model-4088877

Severino, A., Martseniuk, L., Curto, S., & Neduzha, L. (2021). Routes planning models for railway transport systems in relation to passengers' demand. *Sustainability*, *13*(8686). https://doi.org/10.3390/su13168686

Sumito Corporation (2020). *Contributing to the advancement of railway infrastructure development in the growing economy of the Philippines*. http://www.sumitomocorp.com/en/us/business/case/group/26868

Wheeler, J. (2005). Geography. In *Encyclopedia of social measurement* (pp. 115–123). https://doi.org/10.1016/B0-12-369398-5/00277-2

Zhang, C., Yong-sheng, Q., Zeng, J., Guang, X., & Yin, F. (2019). Study on the structure of rail rapid transit trunk line oriented to the development mode of compact urban agglomeration. In *CICTP 2019: Transportation in China—Connecting the world* (pp. 1949–1960). https://doi.org/10.1061/9780784482292.170

APPENDIX A: THE PROCESS FOR SOLVING GREY CORRELATION COEFFICIENT

Table 7. Removing Incommensurability from the Data in Table 3

Year	Population	Perceived Passengers	Annual Income	Number of Tourists
2015	1	1	1	1
2016	1.021699961	1.021700038	1.080957696	1.154191892
2017	1.043400011	1.04339994	1.260453369	1.085353649
2018	1.065099971	1.065099978	1.338627673	1.24756479

Table 8.1. Computing the Difference Between Reference Sequence and Analysis Sequence

Year	Population	Perceived Passengers	Annual Income	Number of Tourists
2015	0	0	0	0
2016	0.0216999607	0.02170003844	0.0809576964	0.1541918919
2017	0.04340001054	0.04339994003	0.2604533693	0.08535364869
2018	0.06509997123	0.06509997847	0.3386276734	0.2475647904

From the data in Table 8.1, the following values were collected:

$Max = 0.3386276734$

$Min = 0$

$Max * \rho = 0.3386276734 * 0.5 = 0.1693138367$

Table 8.2. The Correlation Coefficient

Year	Population	Perceived Passengers	Annual Income	Number of Tourists
2015	1	1	1	1
2016	0.8863958468	0.8863954861	0.6765205559	0.5233719892
2017	0.7959699799	0.7959702437	0.3939663947	0.6648427711
2018	0.7222861068	0.7222860845	0.3333333333	0.4061465993

Table 9. The Grey Correlation Coefficient and Index Weight

	Population	Perceived Passengers	Annual Income	Number of Tourists
Grey correlation	0.8511629834	0.8511629536	0.600955071	0.6485903399
Coefficient index weight	0.2883469105	0.2883469004	0.2035844385	0.2197217505

APPENDIX B: SOLVING FOR THE WEIGHT OF EACH EDGE

To see the reference of each cell in Tables 11 to 41 and how Formula 6 was used, see Table 10 below.

Table 10. Reference for the Formula

	Population	Perceived Passengers	Annual Income	Number of Tourists
City 1	$p1$	$g1$	$e1$	$t1$
City 2	$p2$	$g2$	$e2$	$t2$
2015	P	G	E	T
Parenthesis	$\left(\frac{P}{p_i + p_j}\right)$	$\left(\frac{G}{g_i + g_j}\right)$	$\left(\frac{E}{e_i + e_j}\right)$	$\left(\frac{T}{t_i + t_j}\right)$
Parenthesis * index weight	$0.288\left(\frac{P}{p_i + p_j}\right)$	$0.288\left(\frac{G}{g_i + g_j}\right)$	$0.203\left(\frac{E}{e_i + e_j}\right)$	$0.219\left(\frac{T}{t_i+t_j}\right)$
Distance	$\frac{1}{d_{1,2}}$			
Edge weight	$\frac{SUM(Parenthesis)}{Distance}$			

The following tables below show the actual computation for the weight of each edge.

Table 11. Computing for the Edge Weight of San Jose and Muñoz

	Population	Perceived Passengers	Annual Income	Number of Tourists
San Jose	139,738	90,281	554,412,749	5,197.4
Muñoz	81,483	49,939	389,238,966	5,197.4
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	50.71027163	52.10827985	13.05683672	345.9515335
Parenthesis * index weight	14.62215016	15.02526098	2.658168772	76.01307653
Distance	0.04785			
Edge weight	2,263.712778			

Table 12. Computing for the Edge Weight of San Jose and Palayan

	Population	Perceived Passengers	Annual Income	Number of Tourists
San Jose	139,738	90,281	554,412,749	5,197.4
Palayan	41,041	25,071	272,329,409	5,197.4
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	62.05464683	63.34197066	14.90320318	345.9515335
Parenthesis * index weight	17.8932657	18.26446091	3.034060252	76.01307653
Distance	0.02639			
Edge weight	4,365.474172			

Table 13. Computing for the Edge Weight of Muñoz and Palayan

	Population	Perceived Passengers	Annual Income	Number of Tourists
Muñoz	81,483	49,939	389,238,966	5,197.4
Palayan	41,041	25,071	272,329,409	5,197.4
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	91.55901701	97.40865218	18.62408607	345.9515335
Parenthesis * index weight	26.40075968	28.08748293	3.791574107	76.01307653
Distance	0.01848			
Edge weight	7,266.931453			

Table 14. Computing for the Edge Weight of Muñoz and Cabanatuan

	Population	Perceived Passengers	Annual Income	Number of Tourists
Muñoz	81,483	49,939	389,238,966	5,197.4
Cabanatuan	302,231	199,374	1,056,972,930	5,197.4
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	29.23577717	29.30702771	8.519571989	345.9515335
Parenthesis * index weight	8.430046023	8.450590601	1.73445228	76.01307653
Distance	0.02695			
Edge weight	3,511.249181			

Table 15. Computing for the Edge Weight of Muñoz and Mabalacat

	Population	Perceived Passengers	Annual Income	Number of Tourists
Muñoz	81,483	49,939	389,238,966	5,197.4
Mabalacat	250,799	166,090	999,537,542	500,990
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	33.76101324	33.82241736	8.871914443	7.104279956
Parenthesis * index weight	9.734883862	9.75258921	1.806183721	1.560964828
Distance	0.01172			
Edge weight	1,950.053039			

Table 16. Computing for the Edge Weight of Muñoz and Tarlac

	Population	Perceived Passengers	Annual Income	Number of Tourists
Muñoz	81,483	49,939	389,238,966	5,197.4
Tarlac	342,493	221,706	1,060,385,689	47,696
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	26.45946233	26.89769	8.499514904	67.98763173
Parenthesis * index weight	7.629504216	7.75586554	1.73036897	14.93836146
Distance	0.02053			
Edge weight	1,561.32977			

Table 17. Computing for the Edge Weight of Palayan and Gapan

	Population	Perceived Passengers	Annual Income	Number of Tourists
Palayan	41,041	25,071	272,329,409	5,197.4
Gapan	110,303	71,668	454,728,278	5,197
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	74.12369833	75.52923847	16.946532	345.9515335
Parenthesis * index weight	21.37333941	21.7786218	3.450050203	76.01307653
Distance	0.01984			
Edge weight	6,180.195965			

Table 18. Computing for the Edge Weight of Palayan and Cabanatuan

	Population	Perceived Passengers	Annual Income	Number of Tourists
Palayan	41,041	25,071	272,329,409	5,197.4
Cabanatuan	302,231	199,374	1,056,972,930	5,197
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	32.68013995	32.55418031	9.268851787	345.9515335
Parenthesis * index weight	9.423217389	9.386896988	1.886993987	76.01307653
Distance	0.05051			
Edge weight	1,914.674023			

Table 19. Computing for the Edge Weight of Cabanatuan and Gapan

	Population	Perceived Passengers	Annual Income	Number of Tourists
Cabanatuan	302,231	199,374	1,056,972,930	5,197
Gapan	110,303	71,668	454,728,278	5,197
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	27.19333922	26.95753057	8.150490517	345.9515335
Parenthesis * index weight	7.841115349	7.773120382	1.659313036	76.01307653
Distance	0.03165			
Edge weight	2,947.444717			

Table 20. Computing for the Edge Weight of Cabanatuan and Mabalacat

	Population	Perceived Passengers	Annual Income	Number of Tourists
Cabanatuan	302,231	199,374	1,056,972,930	5,197
Mabalacat	250,799	166,090	999,537,542	500,990
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	20.28493391	19.99272979	5.991268475	7.104279956
Parenthesis * index weight	5.849098023	5.764841666	1.219729029	1.560964828
Distance	0.01361			
Edge weight	1,057.651252			

Table 21. Computing for the Edge Weight of Cabanatuan and Tarlac

	Population	Perceived Passengers	Annual Income	Number of Tourists
Cabanatuan	302,231	199,374	1,056,972,930	5,197
Tarlac	342,493	221,706	1,060,385,689	47,696
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	17.39996805	17.35210174	5.819092831	67.98763173
Parenthesis * index weight	5.01722703	5.003424752	1.184676747	14.93836146
Distance	0.01988			
Edge weight	1,315.074949			

Table 22. Computing for the Edge Weight of Tarlac and Gapan

	Population	Perceived Passengers	Annual Income	Number of Tourists
Tarlac	342,493	221,706	1,060,385,689	47,696
Gapan	110,303	71,668	454,728,278	5,197
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	24.77534475	24.90548924	8.132131726	67.98763173
Parenthesis * index weight	7.143894115	7.181420625	1.655575472	14.93836146
Distance	0.01529			
Edge weight	2,022.187814			

Table 23. Computing for the Edge Weight of Tarlac and Mabalacat

	Population	Perceived Passengers	Annual Income	Number of Tourists
Tarlac	342,493	221,706	1,060,385,689	47,696
Mabalacat	250,799	166,090	999,537,542	500,990
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	18.9083571	18.84140889	5.981342496	6.554016323
Parenthesis * index weight	5.452166352	5.432861851	1.217708254	1.440059939
Distance	0.033			
Edge weight	410.3877696			

Table 24. Computing for the Edge Weight of Gapan and San Jose Del Monte

	Population	Perceived Passengers	Annual Income	Number of Tourists
Gapan	110,303	71,668	454,728,278	5,197
San Jose Del Monte	574,089	371,847	1,078,619,971	27,085
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	16.39144964	16.47435374	8.035425982	111.3961208
Parenthesis * index weight	4.726423862	4.750328838	1.635887687	24.47615068
Distance	0.01335			
Edge weight	2,665.827046			

Table 25. Computing for the Edge Weight of Gapan and Malolos

	Population	Perceived Passengers	Annual Income	Number of Tourists
Gapan	110,303	71,668	454,728,278	5,197
Malolos	252,074	172,135	754,753,646	27,085
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	30.95719927	29.96937281	10.18709426	111.3961208
Parenthesis * index weight	8.926412768	8.641575758	2.073933864	24.47615068
Distance	0.01543			
Edge weight	2,859.239991			

Table 26. Computing for the Edge Weight of Gapan and San Fernando

	Population	Perceived Passengers	Annual Income	Number of Tourists
Gapan	110,303	71,668	454,728,278	5,197
San Fernando	306,659	206,158	1,191,068,312	500,990
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	26.90455485	26.29927725	7.486408974	7.104279956
Parenthesis * index weight	7.75784527	7.583315077	1.524116368	1.560964828
Distance	0.01603			
Edge weight	1,149.484812			

Table 27. Computing for the Edge Weight of Gapan and Angeles

	Population	Perceived Passengers	Annual Income	Number of Tourists
Gapan	110,303	71,668	454,728,278	5,197.4
Angeles	411,634	277,574	1,436,713,978	412,829
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	21.49335456	20.92137544	6.51413297	8.602559551
Parenthesis * index weight	6.197542384	6.032613759	1.326176103	1.890169444
Distance	0.01592			
Edge weight	970.2576439			

Table 28. Computing for the Edge Weight of Gapan and Mabalacat

	Population	Perceived Passengers	Annual Income	Number of Tourists
Gapan	110,303	71,668	454,728,278	5,197.4
Mabalacat	250,799	166,090	999,537,542	500,990
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	31.06650475	30.73134448	8.472389429	7.104279956
Parenthesis * index weight	8.957930666	8.861287925	1.724846645	1.560964828
Distance	0.01129			
Edge weight	1,869.356073			

Table 29. Computing for the Edge Weight of Mabalacat and San Fernando

	Population	Perceived Passengers	Annual Income	Number of Tourists
Mabalacat	250,799	166,090	999,537,542	500,990
San Fernando	306,659	206,158	1,191,068,312	500,990
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	20.12380664	19.62837409	5.624519964	3.588990798
Parenthesis * index weight	5.802637471	5.65978083	1.145064739	0.7885793409
Distance	0.04			
Edge weight	334.9015595			

Table 30. Computing for the Edge Weight of Mabalacat and Angeles

	Population	Perceived Passengers	Annual Income	Number of Tourists
Mabalacat	250,799	166,090	999,537,542	500,990
Angeles	411,634	277,574	1,436,713,978	412,829
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	16.93481001	16.468821	5.057403252	3.935239911
Parenthesis * index weight	4.883100147	4.748733489	1.029608602	0.864657802
Distance	0.11111			
Edge weight	103.7359377			

Table 31. Computing for the Edge Weight of Mabalacat and Olongapo

	Population	Perceived Passengers	Annual Income	Number of Tourists
Mabalacat	250,799	166,090	999,537,542	500,990
Olongapo	233,040	147,005	1,255,249,945	164,042
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	23.18576427	23.33676041	5.464420231	5.407404456
Parenthesis * index weight	6.685543496	6.729082529	1.112470925	1.188124373
Distance	0.013			
Edge weight	1,208.863179			

Table 32. Computing for the Edge Weight of Angeles and San Fernando

	Population	Perceived Passengers	Annual Income	Number of Tourists
Angeles	411,634	277,574	1,436,713,978	412,829
San Fernando	306,659	206,158	1,191,068,312	500,990
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	15.61782866	15.10469227	4.688785067	3.935239911
Parenthesis * index weight	4.503352642	4.355391197	0.9545636753	0.864657802
Distance	0.08			
Edge weight	133.4745665			

Table 33. Computing for the Edge Weight of Angeles and Balanga

	Population	Perceived Passengers	Annual Income	Number of Tourists
Angeles	411,634	277,574	1,436,713,978	412,829
Balanga	96,061	63,979	521,987,321	244,303
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	22.09629207	21.39235492	6.290446821	5.472411936
Parenthesis * index weight	6.371397551	6.168419234	1.280637084	1.20240793
Distance	0.01374			
Edge weight	1,093.366943			

Table 34. Computing for the Edge Weight of Angeles and Olongapo

	Population	Perceived Passengers	Annual Income	Number of Tourists
Angeles	411,634	277,574	1,436,713,978	412,829
Olongapo	233,040	147,005	1,255,249,945	164,042
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	17.40131757	17.20910125	4.576995351	6.233797504
Parenthesis * index weight	5.017616159	4.962191005	0.9318050287	1.3697009
Distance	0.01353			
Edge weight	907.7097629			

Table 35. Computing for the Edge Weight of San Fernando and Malolos

	Population	Perceived Passengers	Annual Income	Number of Tourists
San Fernando	306,659	206,158	1,191,068,312	500,990
Malolos	252,074	172,135	754,753,646	27,084.66667
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	20.07788514	19.31471901	6.332083112	6.809826767
Parenthesis * index weight	5.789396151	5.56933936	1.289113585	1.496267058
Distance	0.03021			
Edge weight	468.1931862			

Table 36. Computing for the Edge Weight of San Fernando and Balanga

	Population	Perceived Passengers	Annual Income	Number of Tourists
San Fernando	306,659	206,158	1,191,068,312	500,990
Balanga	96,061	63,979	521,987,321	244,303
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	27.85602155	27.04784239	7.192472984	4.825078191
Parenthesis * index weight	8.032197754	7.799161516	1.464275574	1.060174627
Distance	0.01783			
Edge weight	1,029.490155			

Table 37. Computing for the Edge Weight of San Fernando and Olongapo

	Population	Perceived Passengers	Annual Income	Number of Tourists
San Fernando	306,659	206,158	1,191,068,312	500,990
Olongapo	233,040	147,005	1,255,249,945	164,042
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	20.78598812	20.6890954	5.036591754	5.407404456
Parenthesis * index weight	5.993575455	5.965636532	1.025371704	1.188124373
Distance	0.01639			
Edge weight	864.7167824			

Table 38. Computing for the Edge Weight of Olongapo and Balanga

	Population	Perceived Passengers	Annual Income	Number of Tourists
Olongapo	233,040	147,005	1,255,249,945	164,042
Balanga	96,061	63,979	521,987,321	244,303
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	34.08733793	34.63117108	6.93273014	8.806516549
Parenthesis * index weight	9.82897858	9.98579084	1.411395973	1.934983232
Distance	0.02101			
Edge weight	1,102.386893			

Table 39. Computing for the Edge Weight of Malolos and San Jose Del Monte

	Population	Perceived Passengers	Annual Income	Number of Tourists
Malolos	252,074	172,135	754,753,646	27,084.66667
San Jose Del Monte	574,089	371,847	1,078,619,971	27,084.66667
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	13.57864852	13.43173671	6.720455801	66.38621482
Parenthesis * index weight	3.915361351	3.872999648	1.368180221	14.58649533
Distance	0.03236			
Edge weight	733.7155919			

Table 40. Computing for the Edge Weight of Malolos and Meycauayan

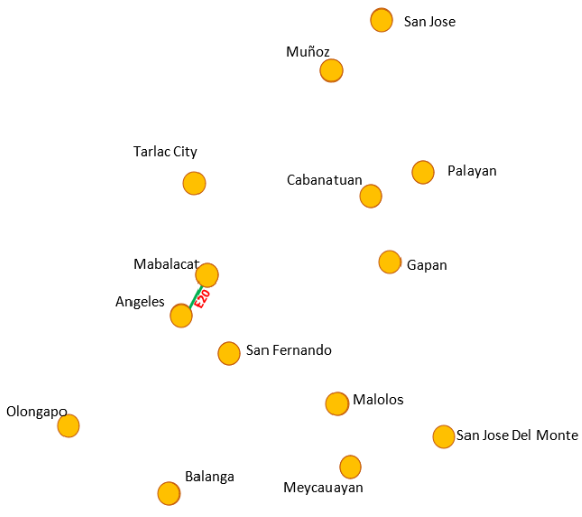
	Population	Perceived Passengers	Annual Income	Number of Tourists
Malolos	252,074	172,135	754,753,646	27,084.66667
Meycauayan	209,083	134,816	907,431,879	27,084.66667
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	24.32615573	23.80387423	7.412593946	66.38621482
Parenthesis * index weight	7.014371851	6.863773353	1.509088777	14.58649533
Distance	0.04425			
Edge weight	677.3724139			

Table 41. Computing for the Edge Weight of San Jose Del Monte and Meycauayan

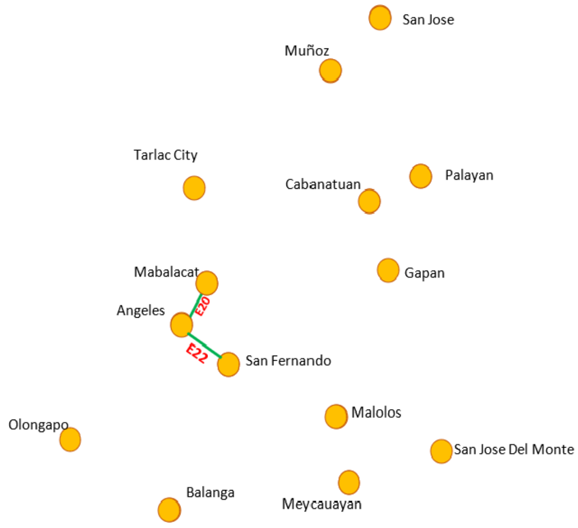
	Population	Perceived Passengers	Annual Income	Number of Tourists
San Jose Del Monte	574,089	371,847	1,078,619,971	27,084.66667
Meycauayan	209,083	134,816	907,431,879	27,084.66667
2015	11,218,177	7,306,623	12,321,106,360	3,596,097
Parenthesis	14.32402716	14.42107081	6.20381908	66.38621482
Parenthesis * index weight	4.130288978	4.158271069	1.263001024	14.58649533
Distance	0.05263			
Edge weight	458.6368308			

APPENDIX C: USING KRUSKAL'S ALGORITHM TO GET THE MINIMUM SPANNING TREE

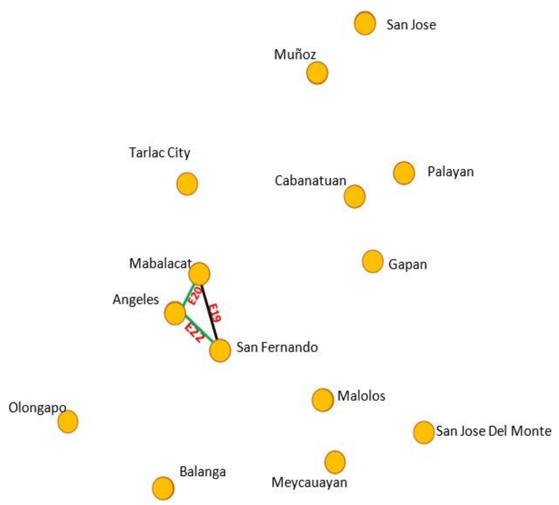
The figure below shows the process of using Kruskal's algorithm to get the minimum spanning tree of the initial railway network. Note that the researchers removed the edges that create a cycle/loop from the figure to highlight the resulting minimum spanning tree.



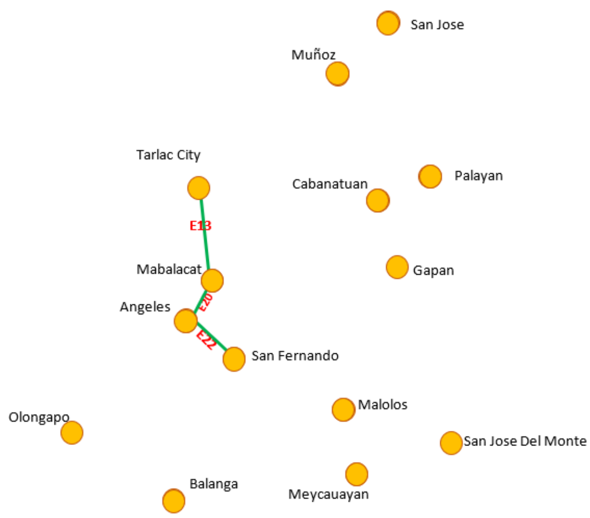
1. Since e_{20} has the lowest weight, the minimum spanning tree will start from here.



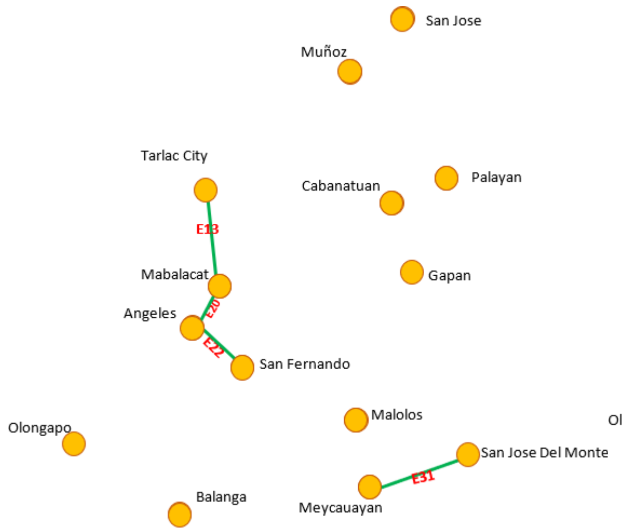
2. The edge e_{22} from Angeles to San Fernando was added



3. The edge e_{19} from Mabalacat to San Fernando was NOT added since it will create a loop

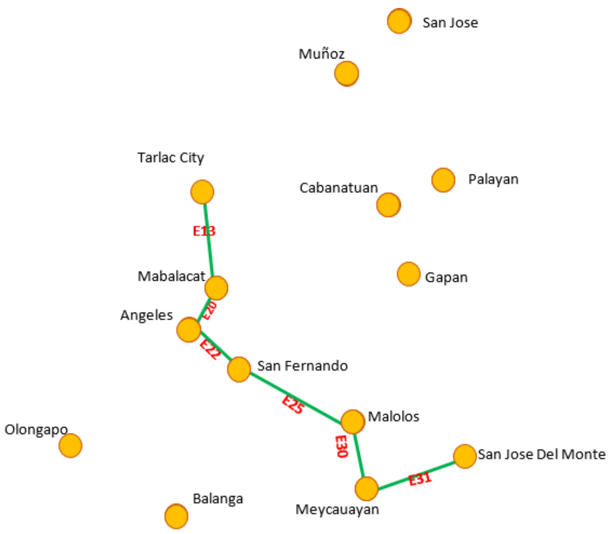


4. The edge e_{13} from Tarlac to Mabalacat was added

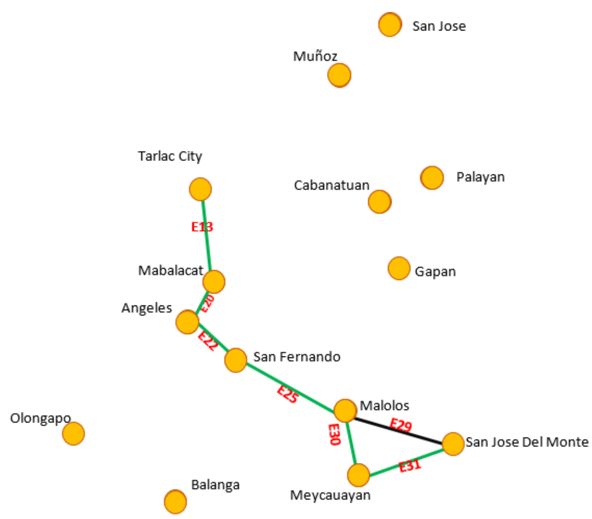


5. The edge e_{31} from San Jose Del Monte to Meycauayan was added

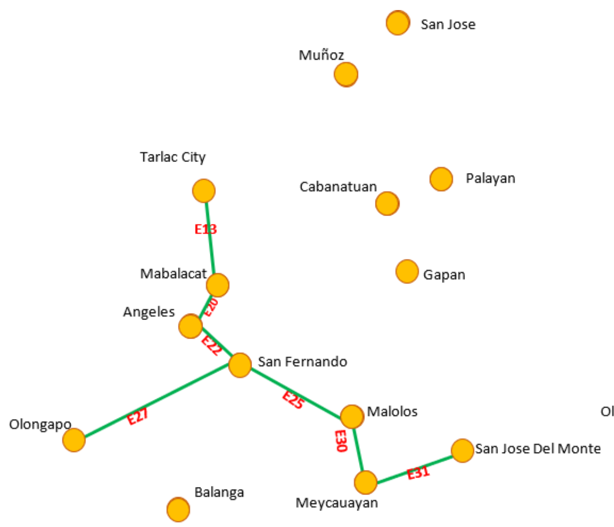
6. The edge e_{25} from San Fernando to Malolos was added



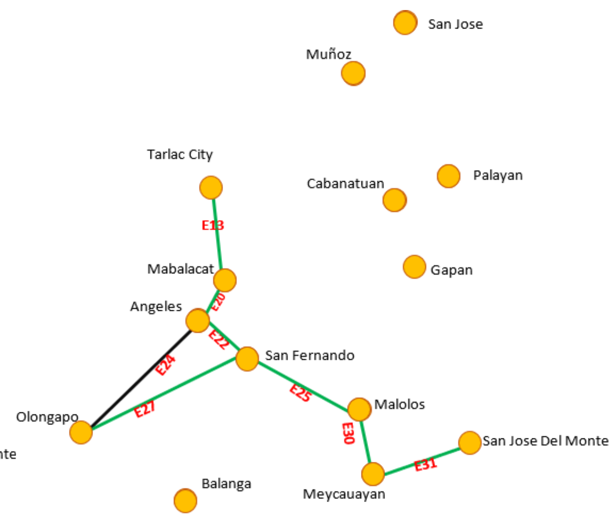
7. The edge e_{30} from Malolos to Meycauayan was added



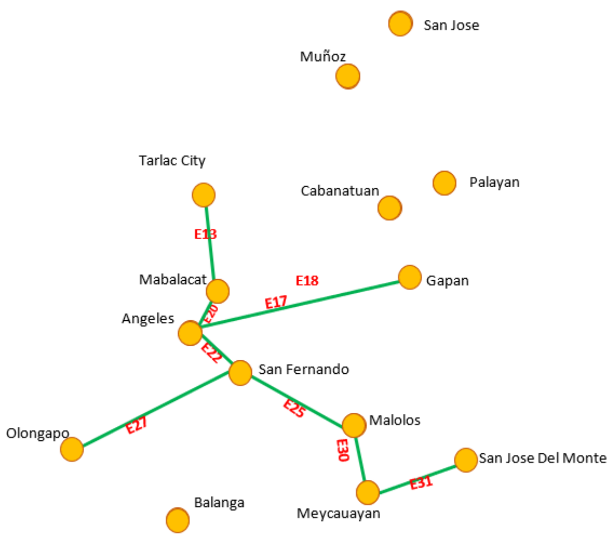
8. The edge e_{29} from Olongapo to Balanga was NOT added since it will create a loop



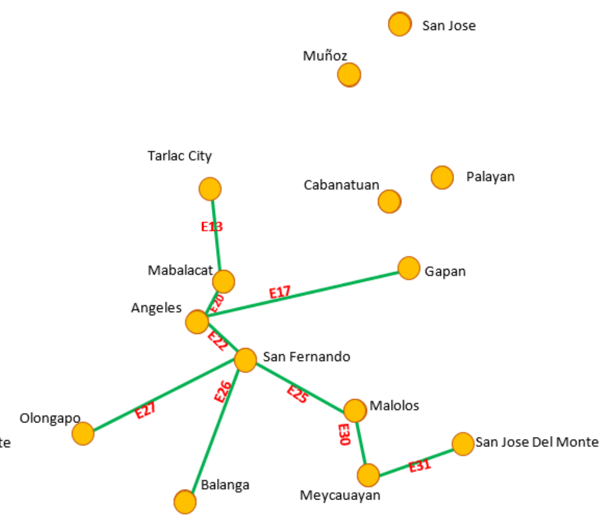
9. The edge e_{27} from San Fernando to Olongapo was added



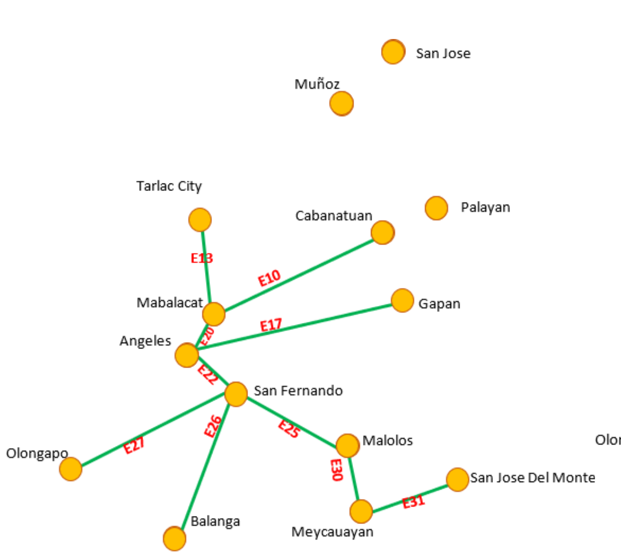
10. The edge e_{24} from Angeles to Olongapo was NOT added since it will create a loop



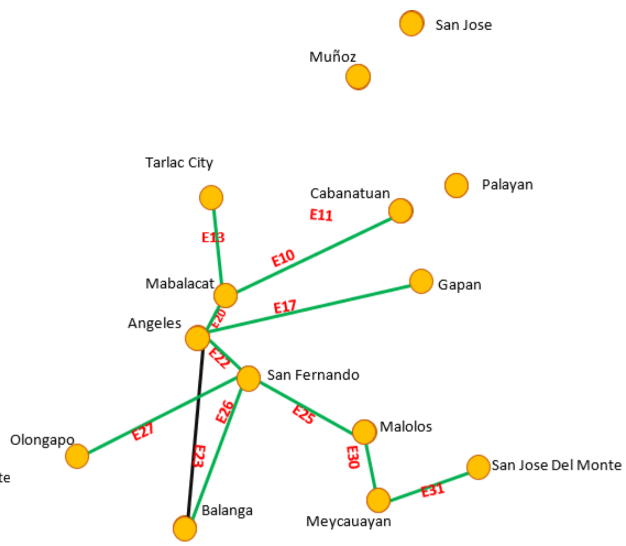
11. The edge e_{17} from Gapan to Angeles was added



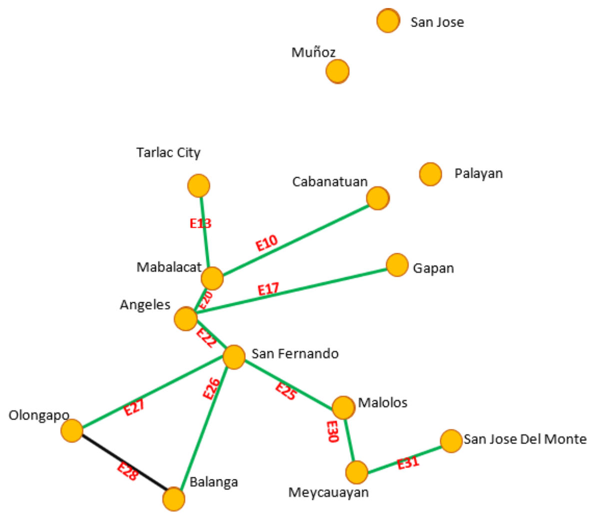
12. The edge e_{26} from San Fernando to Balanga was added



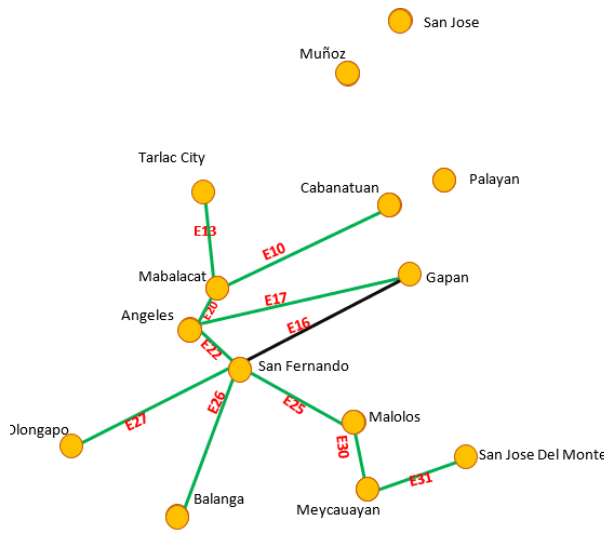
13. The edge e_{10} from Cabanatuan to Mabalacat was added



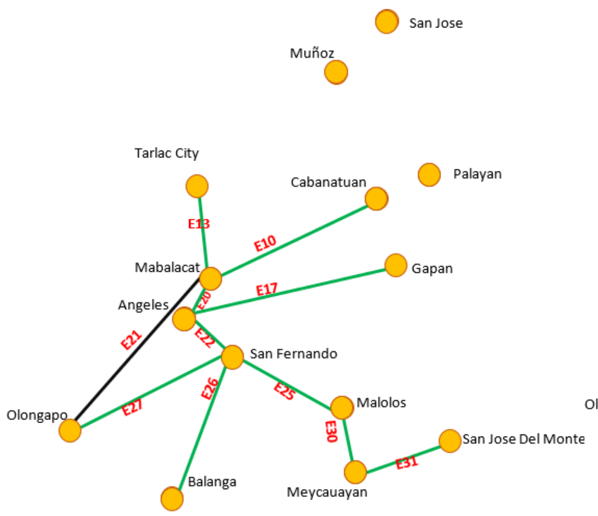
14. The edge e_{23} from Angeles to Balanga was NOT added since it will create a loop.



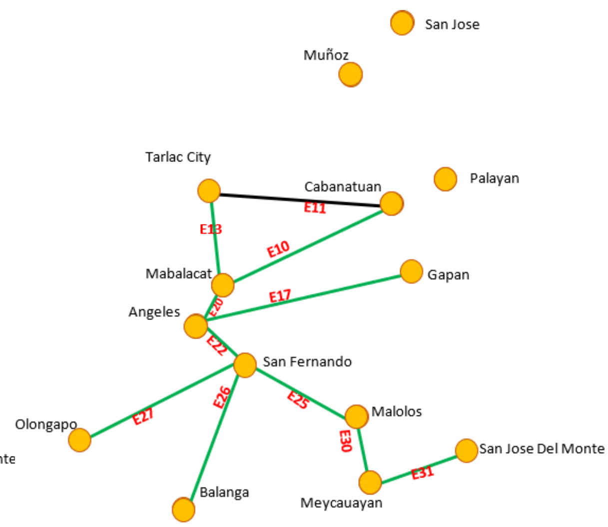
15. The edge e_{28} from Olongapo to Balanga was NOT added since it will create a loop.



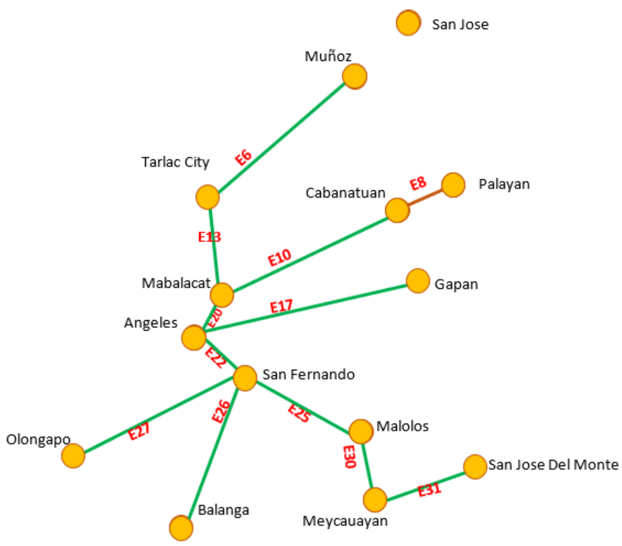
16. The edge e_{16} from Gapan to San Fernando was NOT added since it will create a loop



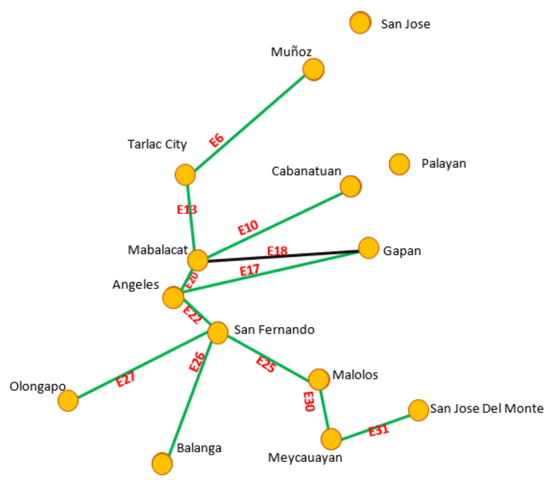
17. The edge e_{21} from Mabalacat to Olongapo was NOT added since it will create a loop



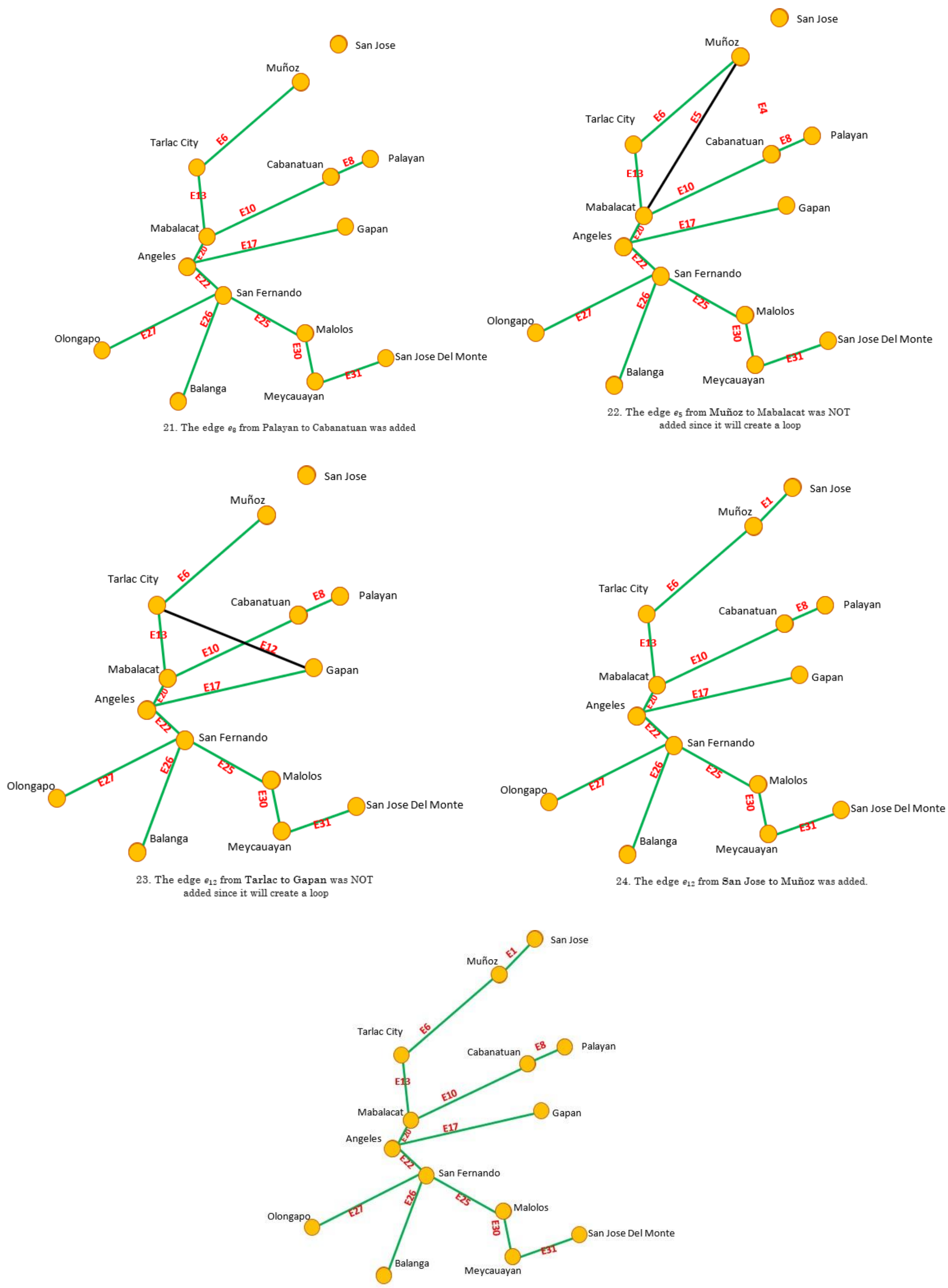
18. The edge e_{11} from Cabanatuan to Tarlac was NOT added since it will create a loop



19. The edge e_6 from Muñoz to Tarlac was added.



20. The edge e_{18} from Gapan to Mabalacat was NOT added since it will create a loop



21. The edge e_8 from Palayan to Cabanatuan was added

22. The edge e_5 from Muñoz to Mabalacat was NOT added since it will create a loop

23. The edge e_{12} from Tarlac to Gapan was NOT added since it will create a loop

24. The edge e_{12} from San Jose to Muñoz was added.

25. Since all the nodes have been passed through, the remaining edges are NOT added, resulting to the minimum spanning tree.

Figure 5. Applying Kruskal's algorithm to the initial railway network.