# An Integrated Approach towards Building a Simulation Model Supporting the Management of the Passenger Transportation System. Part 2 – Case Study.

Joanna Gąbka, Sławomir Susz, Maria Rosienkiewicz
*Wroclaw University of Technology, Poland*

This article presents a simulation model designated as an advising and forecasting tool for designing, redesigning and managing ground-based transportation systems. It considers both public and private transport means. It enables visualisation of the results of changes in the transportation network such as a new transportation mode, schedule adjustment, technology improvements on shuttle speed and other modifications that can influence the effectiveness of the transportation network. The simulation tool enables predictions of future passenger flow size for different means of transport. The simulation tool was developed after thorough analysis of interdependencies between variables in the transportation network model built upon an econometric model, artificial neural network and mathematical model. The simulation model was tested on the real data and determined to be very effective, useful and flexible in use. Successive phases of the model development proved that development of a reliable advising and forecasting tool requires a combination of different methods.

**Keywords:** decision supporting tool, transportation system, algorithm verification, simulation model.

## 1. INTRODUCTION

Development and maintenance of a sustainable passenger transportation system in the contemporary dynamic environment is a challenging task. Organisational units are responsible for the planning and extension of a significant number of communication routes serviced by different means of transport deal with a high level of complexity. Their decision-supporting tools and methods should be flexible and provide the possibility to check the effects and consequences of potential changes to be made in the particular elements of the transportation network and/or schedules. Actual modifications of any part of the system should be done with full awareness and understanding of their positive and negative impacts on the entire system. Decision-making managers of passenger transportation systems also need to be able to forecast passenger demand. It can be accomplished by the use of an appropriate adjustable simulation model. This kind of network concept simulation and analysis tool enables up-front check of possible effects of any contemplated modifications before they are implemented, and also enables passenger demand forecasts and predicted scenarios for the future. To build the best possible and most efficient simulation model supporting management of passenger transportation systems several approaches to generating passenger flows were taken into consideration – econometric model, Artificial Neural Network and Gravity model commonly used in the public transportation field enriched with data processing procedures that enable divide passenger flow within the public transport category into two groups bus and train travellers. The data used to the integrated simulation model verification refer only to people travelling within the county. The results of the study were supposed to support the local authorities in decisions concerning internal transport. For this reason transit passenger flow was extracted from databases.

## 2. METHODOLOGY

The approach proposed in the paper comprises two main stages – initial calculations that are an input for the simulation model (Stage I) and the development of the simulation model (Stage II) (Figure 1). Within stage I three steps were included in order to research and analyze different methods of calculating passenger flow – econometric model (Step I), Artificial Neural Network (Step II), and Gravity model (Step III). Having built the econometric model, the coefficient of determination $R^2$ is computed and verified. If its level is satisfactory, a passenger flow forecast can be implemented in the simulation model. If not, in the step II, the ANN is developed. It was assumed that ANN is the more advanced and complex tool in comparison to the econometric model, so it should be more efficient. And again – having developed the ANN, the coefficient of determination is computed and verified. If its level is acceptable, a passenger flow forecast can be implemented in the simulation model. If not the step III is launched – calculation of the passenger flow based on the Gravity model. In the following calculation it was assumed that the critical value of $R^2$ equals 0,90. The undertaken procedure was inspired by famous William's E. Deming words: *All models are wrong. Some models are useful. Nevertheless, it is important to do everything that is possible so that built models were less wrong and more useful*. The presented research had an experimental character. It was focused on building the most accurate method supporting management of transportation system despite of the obstacles resulting from the data gathering for the purpose of model building. The flexibility was provided by including three different approaches to demand modeling depending on the value of coefficient of determination.

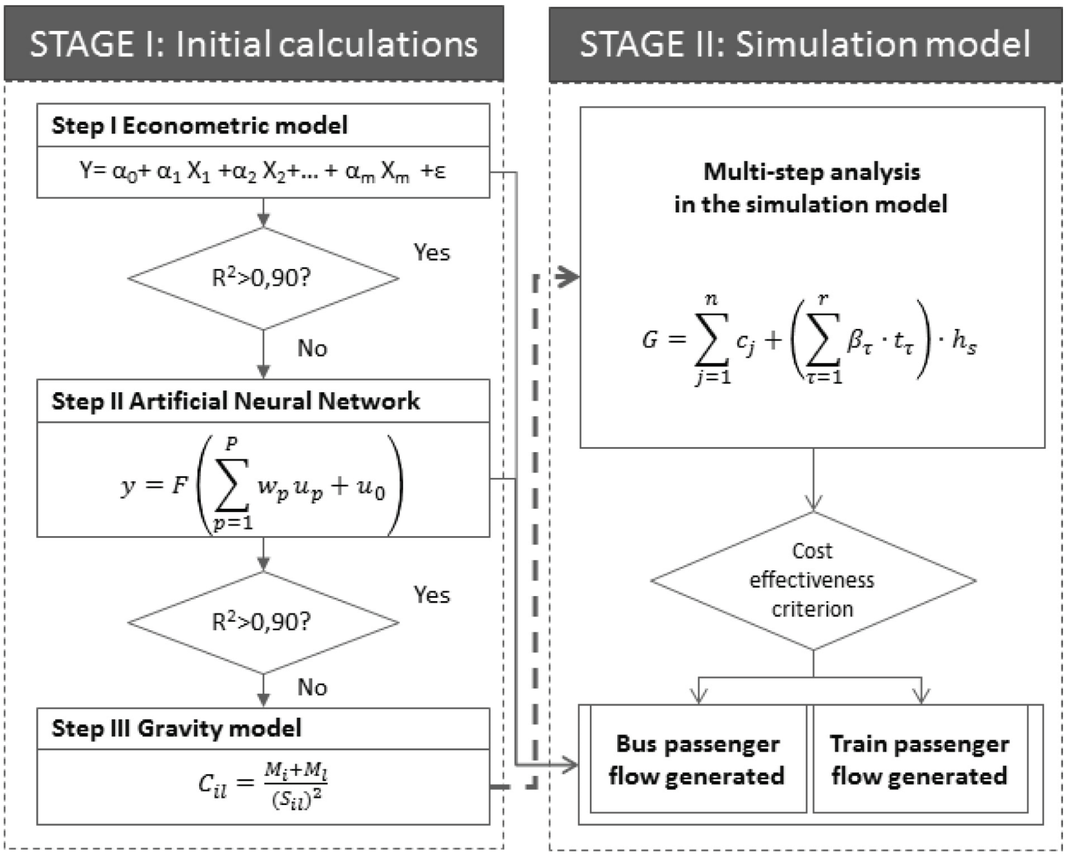



Fig. 1. The methodology used to develop a simulation, advising tool.

Calculations using each method were conducted and the knowledge gained was used to determine which elements were implemented in the embedded process of the decision-aiding simulation tool. The usefulness of the developed hybrid tool was tested on real data from the transportation system of Lower Silesia Region in Poland.

## 3. STAGE I - STEP I: ECONOMETRIC MODELLING

The first step of the Stage I taken into reflection was econometric modelling. The idea was to build two econometric models – the first, in which dependent variable is the value of the bus passenger flow labelled $\hat{y}_b$ (model 1), and the second, in which dependent variable is the value of train passenger labelled $\hat{y}_t$ (model 2). Firstly, on the basis of literature analysis and experts talks, the set of potential explanatory variables was defined. Thirty variables were included in the set. These variables were characterizing transportation points, respectively A and B. The set of the potential variables included: GDP: point A ($x_1$), B ($x_2$); average income *per capita* A ($x_3$), B ($x_4$), unemployment rate A ($x_5$), B ($x_6$); bus's shuttle frequency between A and B ($x_7$); trains shuttle frequency between A and B ($x_7$); ticket price ($x_8$); fuel price ($x_9$); transit time between A and B ($x_{10}$); number of inhabitants A ($x_{11}$), B ($x_{12}$); birth rate A ($x_{13}$), B ($x_{14}$); number of employed A ($x_{15}$), B ($x_{16}$); number of students A ($x_{17}$), B ($x_{18}$); number of people over 60 years old A ($x_{19}$), B ($x_{20}$); number of visiting tourists A ($x_{21}$), B ($x_{22}$); distance from the border (Cz) A ($x_{23}$), B ($x_{24}$); distance from the border (G) A ($x_{25}$), B ($x_{26}$)distance between A and B ($x_{27}$); technical velocity ($x_{28}$) and population density A ($x_{29}$), B ($x_{30}$). Theoretically, this approach looked very promising – unfortunately gathering all the needed data was, in fact, hard (or in some cases impossible) to accomplish. In the researched region gathered data were incomplete, which caused that many observations had to be removed from the sample. Consequently, the sample size shrined, which always negatively influences the quality of the model. Having gathered all the available data and having conducted all the necessary calculations, including variable selection based on the information criteria, econometric model (model 1) of the bus flow was built. It should be underlined, that some of the variables were removed from the model due to the lack of statistical significance. The one-equation linear econometric model is presented below:

$$\hat{y}_b = -265{,}387 - 0{,}029x_3 - 0{,}030x_4 + 13{,}650x_7 - 0{,}009x_{11} + 0{,}126x_{13} + 0{,}101x_{14} + 0{,}003x_{16} - 0{,}007x_{18} + 0{,}046x_{19} \quad (1)$$

where:

$x_3$ – average income *per capita* in the point A,
$x_4$ – average income *per capita* in the point B,
$x_7$ – bus's shuttle frequency between A and B points,
$x_{11}$ – number of inhabitants in the point A,
$x_{13}$ – birth rate in the point A,
$x_{14}$ – birth rate in the point B,
$x_{16}$ – number of employed in the point B,
$x_{18}$ – number of students in the point B,
$x_{19}$ – number of people over 60 in the point A.

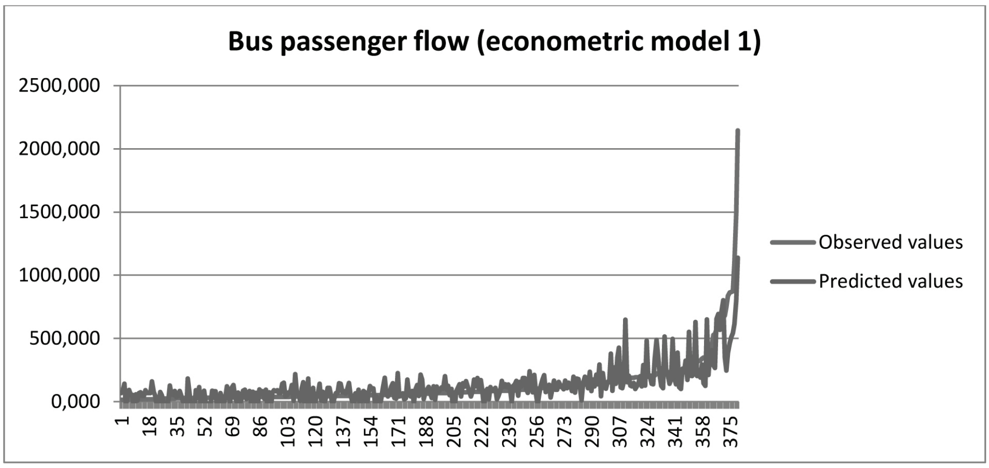



Fig. 2. Plot of observed vs. predicted values – bus passenger flow (econometric model).

Table 1. Statistical summary (bus passenger flow model).

| | b | Stand. Error | t(370) | p-value |
|---|---|---|---|---|
| **Intercept** | 265,3866 | 39,89630 | 6,65191 | 0,000000 |
| **x3** | -0,0298 | 0,00739 | -4,02671 | 0,000069 |
| **x4** | -0,0303 | 0,00807 | -3,75905 | 0,000198 |
| **x7** | 13,6498 | 0,60139 | 22,69724 | 0,000000 |
| **x11** | -0,0091 | 0,00169 | -5,36252 | 0,000000 |
| **x13** | 0,1261 | 0,03204 | 3,93462 | 0,000100 |
| **x14** | 0,1008 | 0,03691 | 2,73082 | 0,006620 |
| **x16** | 0,0027 | 0,00109 | 2,50943 | 0,012519 |
| **x18** | -0,0072 | 0,00330 | -2,18725 | 0,029349 |
| **x19** | 0,0460 | 0,00839 | 5,48683 | 0,000000 |

Coefficient of determination $R^2$ of the model equals 0,611, which means that the model is not properly fitted to the empirical data. Only in 61% it explains the variability of the flow. Although all the variables in the models are statistically significant, the relations and impact of the selected variables seems hard to accept. The econometric model (model 2) of the train flow was built adequately:

$$\hat{y}_t = -1021{,}433 + 117{,}238x_7 + 0{,}001x_{12} \quad (2)$$

where:
$x_7$ – bus's shuttle frequency between A and B points,
$x_{11}$ – number of inhabitants in the point A.

Having analyzed the structure of the model it can be noticed, that only two variables were selected by the information criteria.

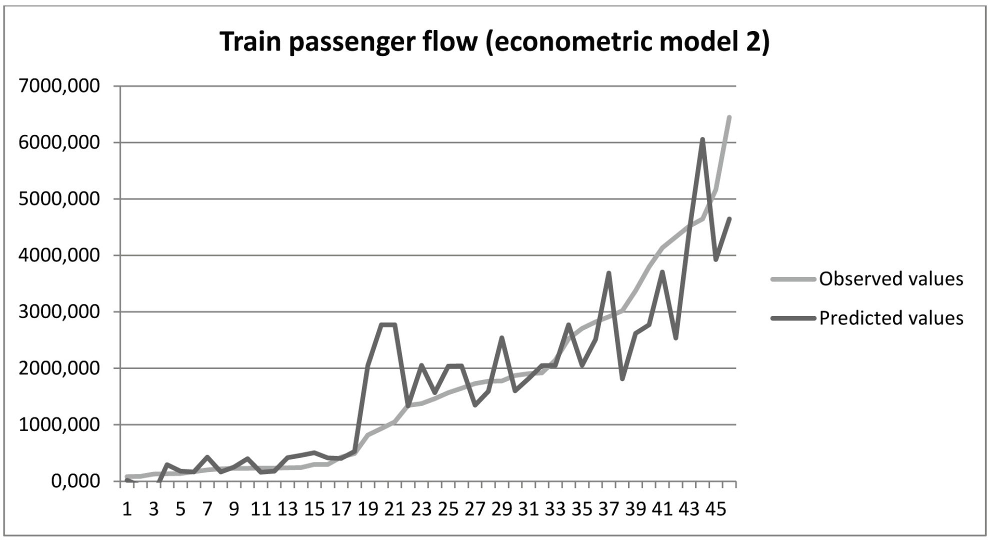



Fig. 3. Plot of observed vs. predicted values – train passenger flow (econometric model).

Table 2. Statistical summary (train passenger flow model).

| | b | Stand. Error | t(43) | p-value |
|---|---|---|---|---|
| **Intercept** | -1021,4334 | 250,5290 | -4,07711 | 0,000193 |
| **x7** | 117,2378 | 11,7046 | 10,01640 | 0,000000 |
| **x12** | 0,0012 | 0,0005 | 2,20253 | 0,033037 |

Coefficient of determination $R^2$ of the model equals 0,795, which means that the model quite well fits the empirical data. In 79,5% it explains the variability of the train passenger flow.

To verify whether linear models are the best possible solution in the analysed aspect, also nonlinear models were explored. Calculations were conducted for i.e. power models and logarithm models. In the paper only results for the best of the analysed model is presented. Nevertheless, even the best of the analysed models is characterised by a very low coefficient of determination (24%). The highest rate of the $R^2$ was achieved for the following example:

$$\hat{y} = 3033 + 28815{,}9x_1 \text{-} 36053{,}7x_2, \quad (3)$$

where:
$x_1$ – ln (sum of the inhabitants in points A and B) / distance,
$x_2$ – ln (average income in the A and B points) / distance.

This analysis indicates that passenger flow is mostly dependent on the number of inhabitants in the transportation points, average income of inhabitants in the transportation points and frequency of the shuttles, however, the coefficient of determination rate is quite low. This result induced further search of models more precisely describing the relations existing in the transportation network. According to the proposed methodology, due to the reason that $R^2$ for both model is <0,90, ANN had to be developed.

## 4. STAGE I - STEP II: ARTIFICIAL NEURAL NETWORKS

According to the second step of the Stage I, two types of ANN were developed – the first one where dependent variable is the value of the bus passenger flow ($ANN_{bus}$), and the second, in which dependent variable is the value of train passenger ($ANN_{train}$). There were the same sets of potential explanatory variables used as in case of econometric models. All together four networks were developed – two comprising of 30 explanatory variables (full potential explanatory variables set) and two consisting of variables selected by the information criteria (ANN$_{bus}$: $x_3$, $x_4$, $x_7$, $x_{11}$, $x_{13}$, $x_{14}$, $x_{16}$, $x_{18}$, $x_19$; ANN$_{train}$: $x_7$, $x_{11}$ ). The first type of ANN performed better in the case of forecasting bus passenger flow, whereas the second performed better in the case of forecasting train passenger flow. Therefore, from the four ANN, two were chosen – characterized by the highest Quality level. The first ANN$_{train}$ based on the whole potential explanatory variable set (MLP 30-13-1) and the second ANN$_{bus}$ based on explanatory variables selected by the information criterion (MLP 2-6-1). Obtained results show that there is no one optimum approach to ANN construction. The available data and its quality defines the way of ANN building. It is one of the reasons why ANN was acknowledged as not an optimal tool that could be input to the simulation model.

Table 3. Artificial Neural Networks of the bus flow and the train flow.

| | ANN (bus) | ANN (train) |
|---|---|---|
| **ANN name** | MLP 30-13-1 | MLP 2-6-1 |
| **Quality (learning)** | 0,866 | 0,910 |
| **Quality (testing)** | 0,977 | 0,577 |
| **Quality (validation)** | 0,947 | 0,854 |
| **Error (learning)** | 2254,4 | 244683,8 |
| **Error (testing)** | 1488,2 | 555752,8 |
| **Error (validation)** | 7569 | 624999,8 |
| **Learning algorithm** | BFGS 21 | BFGS 6 |
| **Error function** | SOS | SOS |
| **Activation (hidden)** | Tanh | Tanh |
| **Activation (output)** | Exponential | Exponential |

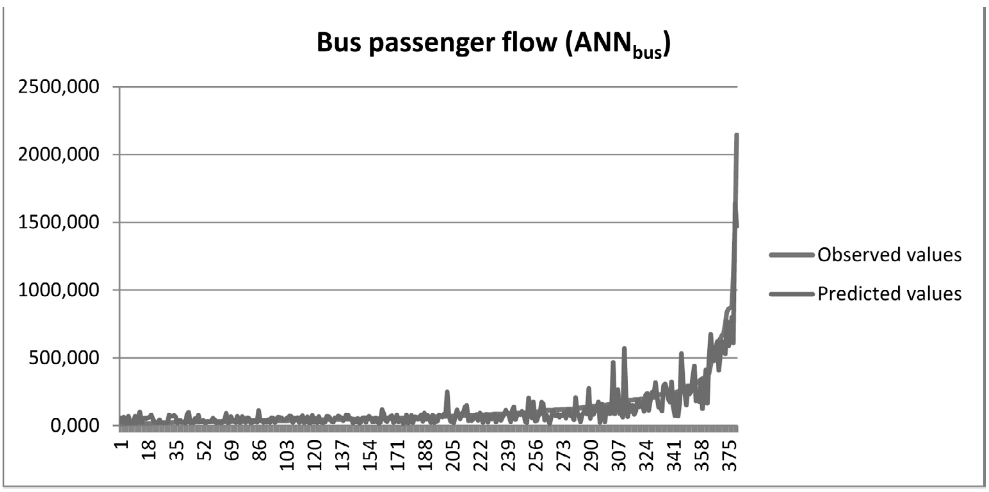



Fig. 4. Plot of observed vs. predicted values – bus passenger flow (ANN$_{bus}$).

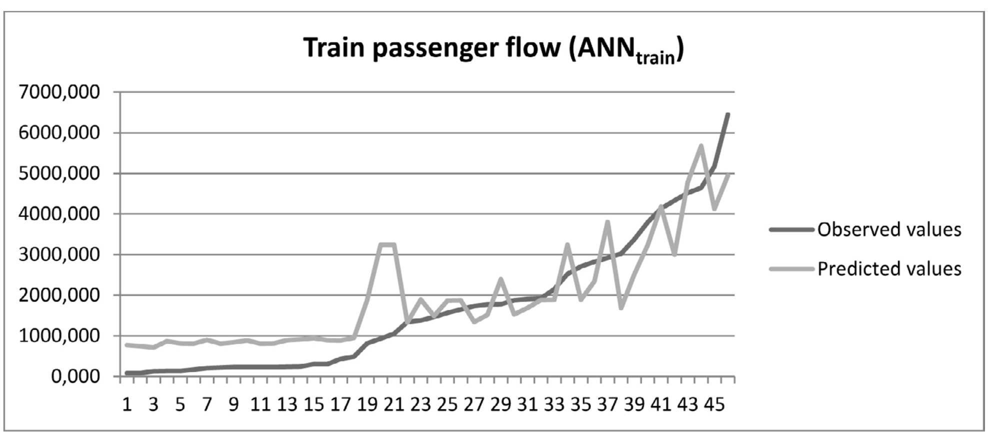



Fig. 5. Plot of observed vs. predicted values – train passenger flow (ANN$_{train}$).

Moreover, their quality was not significantly better to form the econometric models. The coefficient of determination of the $ANN_{bus}$, in which dependent variable was the bus flow, equals 84%, whereas for the $ANN_{train}$ equals 74%. The obtained results are presented in Table 3, Figure 4, 5 and 6.

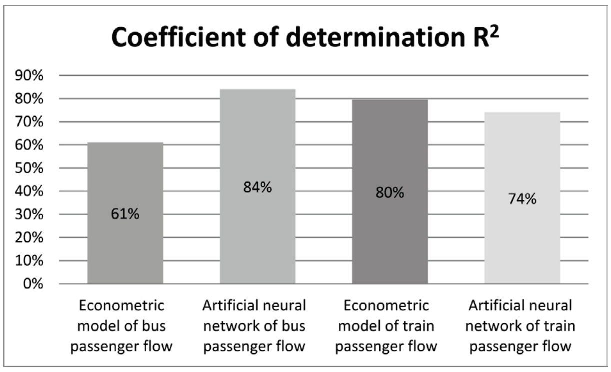



Fig. 6. Coefficient of determination - econometric models and ANN.

The obtained results show that analysed approaches, both econometric modelling and ANN, are not entirely satisfactory in the case of the Lower Silesia region. In both cases – econometric modelling and ANN – the coefficients of determination were lower than 0,90. This is why step III of the proposed methodology was launched.

## 5. STAGE I - STEP III: GRAVITY MODEL

The enriched Gravity model enabled to precisely calculate what is the demand for each transportation mode within the public transport category. After implementing data processing procedures, described by mathematical notation in section 5, to the simulation model the map presenting current situation could be generated. Figure 7 and Figure 8 presents the results.

The performed calculations clearly show that in case of data gathered in the Lower Silesia region econometric models and Artificial Neural Networks are not efficient tools of passenger flow forecasting ($R^2$ too low). Obtained results don’t prove however that perhaps in case of data from other region, these approaches could be useful and more efficient.

Nevertheless, one of the aims of building a simulation model supporting the management of passenger transportation system was to create a tool that can be flexible and used in different

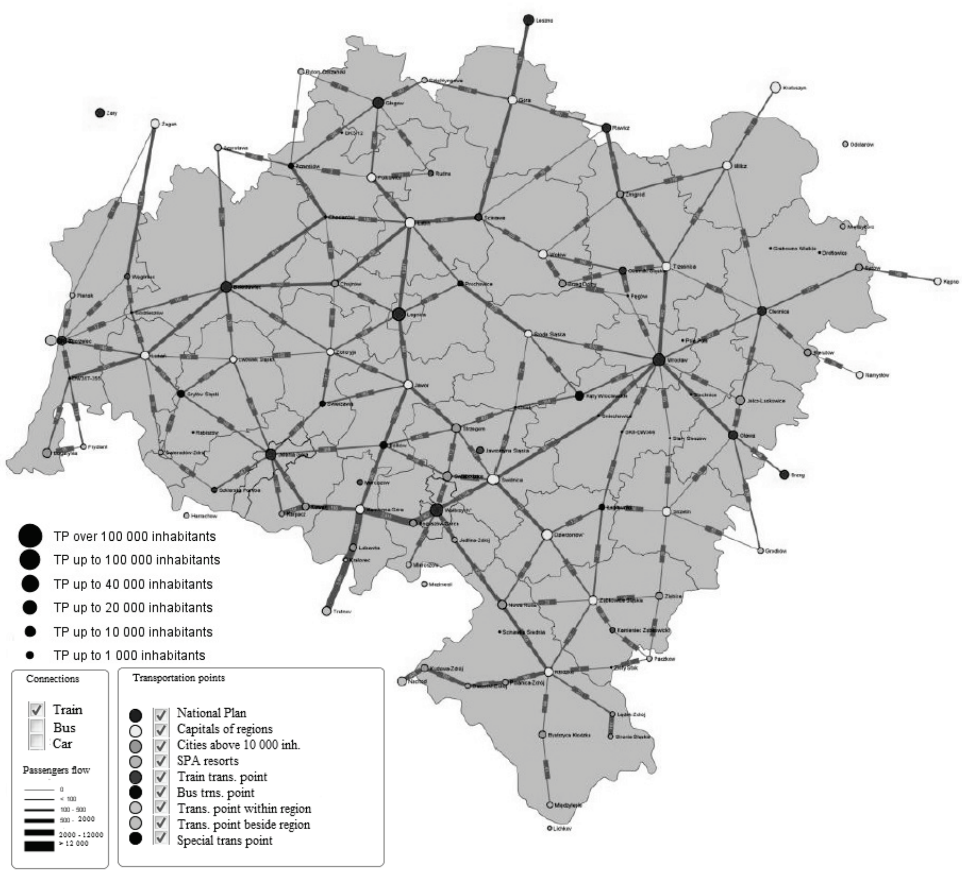



Fig. 7. The demand for bus transportation in the Lower Silesia region (2013).

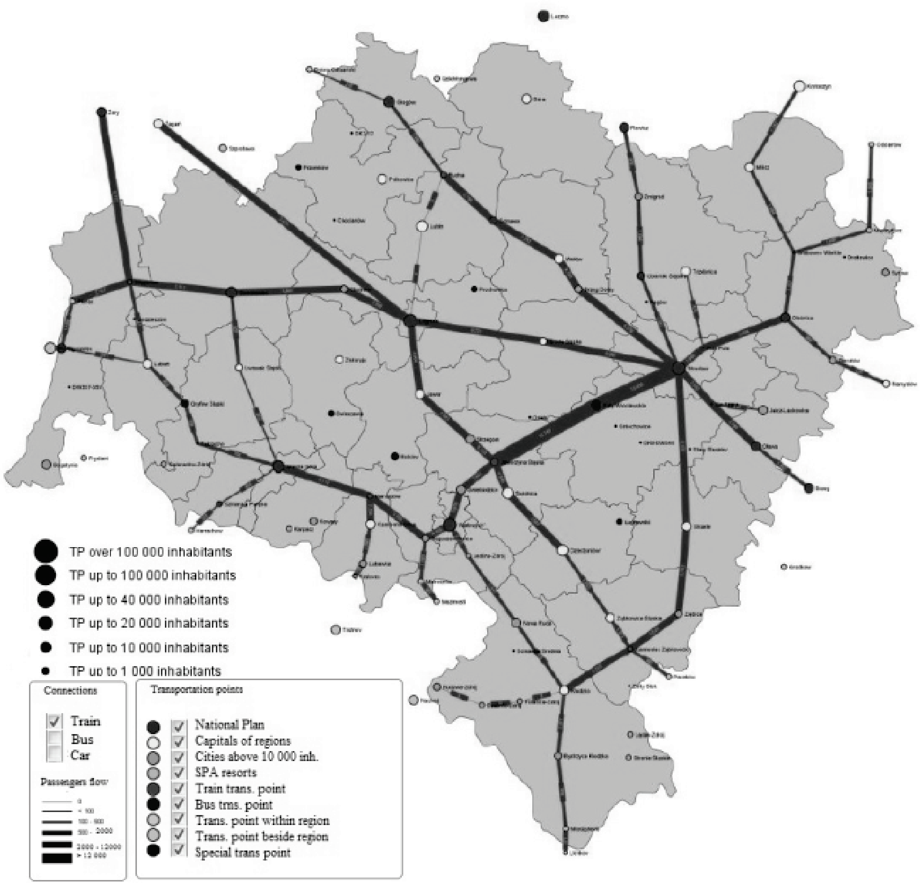



Fig. 8. The demand for railway transportation in the Lower Silesia region (2013).

environments. Implementing econometric modelling and ANN preclude the simulation model from being flexible and versatile. The initial idea of using these methods was to build one econometric model or one ANN that could, later on, be implemented to the simulation model. Unfortunately, the obtained results show that such an approach would require building a new model or a new ANN every time the different application is necessary. It means that whenever there is a need to use the simulation model supporting management of passenger flow in a different transportation system, the new set of explanatory variables should be defined, a new large set of data should be collected and new econometric model or ANN should be build.

## 6. STAGE II: DESCRIPTION OF THE PROPOSED SIMULATION TOOL

The traditional approach to analysing public transport demand is essentially static, with the emphasis on equilibrium states. Nowadays greater emphasis is placed on the dynamic processes that allow these equilibrium states to be achieved or in certain circumstances prevent equilibrium states from being achieved, [1], [4], [7]. The usefulness of the simulation model in practice is highly dependable on implemented algorithms and data processing procedures altogether with data basis containing reliable and complete data set. Taking into consideration results obtained from the econometric model and Neural networks and architecture of the simulation model was built and described in appropriate formulas and equations.

The simulation model built is dedicated to predicting accurately an amount of passengers travelling on a particular route by a specific mode of transport. The model pictures dependencies between factors that influence an amount of travellers in the flow and shows which alternative scenario of the route would be chosen in the given circumstances. It gives the opportunity to plan a properly number of the public transport means and predict a load of roads considering car traffic. The developed instrument supporting passenger transportation system management was built on a base of standard elements necessary in logistic systems and procedures that are common for demand forecasting. The advantage of the proposed simulation model is its adaptability and flexibility. It enables to implement any configuration of transport points within the transportation system that is to be simulated.

*6.1. GENERAL INFORMATION ABOUT VARIABLES*

The **transportation point** is defined as a place which generates 'in' and 'out' passenger flow and directs the flow into different ways. The transportation point in most cases is a city, a town or a village as a source or an aim of the passenger flow. Each point is characterised by an amount of citizens, and they average income. The model also includes **special transportation points** represented by main crossroads which distribute traffic. In this context both characteristics (number of citizens and average earnings) equal 0. It implies that the special point does not have the ability to generate and receive traffic. Although, it is necessary to picture appropriately travelling routes in the model.

**Mode of transport** – possible way of travelling between two transportation points. The exemplary and the most popular transport modes are a railway, a bus and a car. They can be divided into two categories: public organised, subsided by local authorities like railway, bus and private like cars. The simulation model enables to insert any means of transport as far as routes and transport points for them are known.

**Connection** - course for travel from one place to another (direct between two neighbouring transportation points). Each connection is described by length (given in kilometres [km]) and travel time (given in minutes [min.]) and a type of available transport modes (railway, bus, car). The connections are divided into two categories (public, private).

**Transportation point potential** – the potential of a transportation point $i$ towards a transportation point $l$ specified as $C_{il}$. It is the ability of the given point to generate and receive specified number of travelers. The potential is calculated as a ratio of total population living in the point $i$ multiplied by the total population living in the point $l$ and the square of distance between these two points. The formula for $C_{il}$ calculation is presented in (7).

**Transportation network** – is defined as a set of transportation points and their direct connections enabling passengers transport.

**Route** – it is an ordered set of connections occurring in the transportation network which links two randomly chosen transportation points. The routes are divided into two groups: public (public connections: railway, bus) private connections: car.

**Route designating** – it is a procedure of choosing an optimal route between two points. Optimal is the one representing the highest customer satisfaction (9), (10). This criterion will be thoroughly described in the further part of the article.

**Passengers potential** – it is a maximal number of passengers who want to travel between two transportation points. This number is calculated in the mathematical model built for this need.

**Maximal passenger demand** (for chosen connection) – the sum of the passengers potentials for all routes which use this connection.

It is recommended to build simulation model using a mathematical notation as a base describing all relations in the form of mathematical equations. This approach enables to avoid errors which may occur in the case of approximation made for numerical models. The solution implemented in the proposed simulation tool operates on the foundation of the multistage model based on the classic model but enriched with additional formulas. The following steps are distinguished:

1. Determining a maximal number of passengers who want to move from each transportation point in a given period of time.
2. Calculating passengers potentials between each transportation point and any other transportation point (determining where passengers from each point want to travel).
3. Route designation in the transportation network which means indicating optimal public and private route between each pair of transportation points generating and receiving passengers flow.
4. Fixing how the calculated passenger potentials will spread between private and public routes.
5. Calculating maximal passenger demand for each connection in the network. This operation is summarising all potentials resulting from all routes that use given connection.
6. Forecasting.

The listed steps are thoroughly explained in the subsequent subsection.

*6.2. MATHEMATICAL NOTATION OF RELATIONS AND CONDITIONS CONTAINED IN THE SIMULATION TOOL*

The transportation plan suggested by the system is focused on maximisation of the customer/traveller satisfaction. It is defined as a factor that directly influences which kind of the transportation mode and which connection will be chosen by a traveller. The approach chosen to be used in the model assumes that the primary factor motivating traveller to choose a given way of travelling is the generalised cost of the journey between two transportation points as it is explained in [11]. The satisfaction criterion, therefore, is mainly dependent on the cost of the trip, a time and a value of the time for each passenger.

The data necessary for the appropriate model functioning are acquired from the formulas presented below. They are described by the order of the analysis steps presented in the previous section.

1. Determining a maximal number of passengers who want to move from each transportation point in a given period of time. The proposed simulation tool uses Daly's mathematical trip distribution model. It enables to calculate the probability that a given inhabitant of a transportation point undertakes travel, [3]. The following formula enables to calculate probability $\eta_i$ that a single inhabitant of a transportation point $i$ decides to begin a travel.

$$\eta_i = \frac{b}{1 + e^{(-\lambda_i)}} \tag{4}$$

$$\lambda_i = \sum_m a_m \cdot x_m \tag{5}$$

Where:
- $\lambda$ – utility of the travel – dependents on such a parameters as the inhabitant income, motorization level of the region, wealth of the inhabitants comprised in $x_m$,
- $b$ – parameter which value is estimated to correlate model with the real data gathered from the past,
- $m$ – index labeling parameters combined in variable $x$,
- $a_m$ - correlation parameter which indicates power of the influence on the probability. In this case it was fixed that $a_m = 1$,
- $x_m$– average income of inhabitant.

A product of probability value ($\eta_i$) and number of inhabitants of a given transportation point gives figure indicating how many people is to be moved beside the given point in the given time period.

$$P_i = M_i \cdot \eta_i \tag{6}$$

Where:
- $P_i$ - transportation point potential for point $i$.
- $M_i$ – number of inhabitants for point $i$.

2. Calculating passengers potentials between each transportation point and any other transportation point (determining where passengers from each point want to travel). This part of the data processing in the simulation tool relies on the very popular concept known as a gravity model, [10]. In this phase, the model has to find out the destination of travel for the passengers from each point $i$. This phase shows how big the passenger flow is depending on the direction of the passage.

$$C_{il} = \frac{M_i + M_l}{(S_{il})^2} \tag{7}$$

Where:
- $C_{il}$ - gravity of a point $i$ towards point $l$. The parameter which indicates the willingness of the people to move from point $i$ to point $l$.
- $M_i$ - number of inhabitants for point $i$.
- $M_l$ - number of inhabitants for point $l$.
- $S_{il}$ - distance between point $i$ and $l$ given in kilometers.

The model determines passenger potential between point $i$ and every other point dividing the total potential proportionally to each point share in a total gravity.

$$P_{il} = P_i \cdot \frac{c_{il}}{\sum_{j=1}^{k} c_{jl}} \tag{8}$$

3. Route designation in the transportation network which means indicating optimal public and private route between each pair of transportation points generating and receiving passengers flow. On this stage, the model fixes two types of routes (public, private) between each pair of points entered in the transportation network. The simulation tool

conducts analysis of every possible combination of connections that lead from a given point $i$ to point $l$. The brute force algorithm searches for the optimal solution in the sequences on the base of implemented conditions and rules. The rout is optimal for a passenger taking into consideration the satisfaction criterion. It is expressed in function minimizing cost and time.

The satisfaction criterion on this stage is calculated separately for the two categories of routes (public, private) as:

$$S1 = \min_{R_{<1,a>}} (costR1 * timeR1) \quad (9)$$

$$S2 = \min_{R_{<1,a>}} (costR2 * timeR2) \quad (10)$$

Where:
- cost$R1$ – cost of a travel in case of public routs,
- cost$R2$ - cost of a travel in case of private routs,

Time of the travelling for the public routes (rail, bus) is calculated in the following way:

$$time\ R1 = t_{DS} + t_p + t_{ch} + t_{SD} \quad (11)$$

Where:
- $time\ R1$- time of a travel in case of public routs,
- $t_{DS}$ – time 'door – bus stop' – average time of getting to the bus stop/railway station and a wait time for the bus/rail departure,
- $t_p$ – specific travel time between the initial and final bus stop/railway station,
- $t_{ch}$ - total of the change/ transfer time a bus to a train or train to a bus,
- $t_{SD}$ – time 'bus stop –door' - average time of getting from the final bus stop/railway station to the travel destination.

Time of the travelling for the private route is calculated with the following formula:

$$time\ R2 = t_{DC} + t_c + t_{CD} \quad (12)$$

Where:
- $time\ R2$ - time of a travel in case of private routs,
- $t_{DC}$ - time 'door – car' – average time of getting to the car,
- $t_c$ - specific travel time for a car,
- $t_{CD}$ – time 'car-door' - average time of parking a car and getting to the travel destination.

4. Fixing how the calculated passenger potentials will spread between private and public routes.

   The algorithm implemented in the proposed simulation tool specifies the size of passenger flow for each group of routes (public, private) on the base of generalised cost, [5].

$$G = \sum_{j=1}^{n} c_j + \left( \sum_{\tau=1}^{r} \beta_\tau \cdot t_\tau \right) \cdot h_s \quad (13)$$

Where:
- $G$ - size of the passengers flow;
- $J = \{1, \dots, n\}$ - set of a travel cost items $j, n \in J$;
- $j$ – index labelling a given travel cost item,
- $n$ - total number of travel cost items,
- $c_j$ – travel cost items for a given rout,
- $L = \{1, \dots, r\}$ - set of a travel time items $o, r \in L$;
- $o\tau$ - index labeling a given travel time item,
- $r$ - total number of travel time items,
- $\beta_j$ – ratio_for the travel time (the ration equals 1 in the model)
- $t_j$ – travel time items for a given rout,
- $h_s$ – unit cost of a given travel time $s$.

Calculating a unit cost of a travel time for each traveller is of key importance for the proper model performance. The formula determining how the passengers' flow spreads between two kinds of routes is a linear function. In the case when the travel time for the two kinds of routes is different there are two different results for the generalised cost. These functions intersect at point $h_0$ (Figure 9). This point fixes neutral travel cost. If the travel time for public route is longer than the travel time for private route (which is true in most cases) then passengers for who the value of unit cost ($h_s$) is smaller than $h_0$ (people who have smaller income) choose public transport. Rest of the passengers choose private transport. In the Figure 6 $h_{min}$ labels the lowest unit cost of a given travel time and $h_{max}$ marks the highest unit cost of a given travel time (this value is characteristic for the inhabitants revealing the highest income, their time has the highest value).

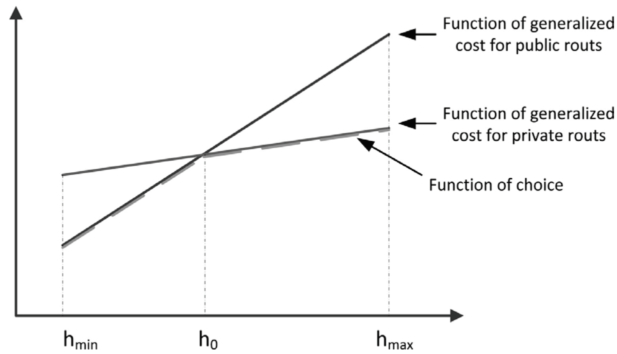


Fig. 9. Generalised cost functions, [2].

The algorithm calculates number of passengers for who the $h_s < h_0$ and the number of rest. As in the [6, 8, 9] distribution of $h_s$ value is assumed as identical to the income distribution in the society. This distribution is approximated with the log-normal or Pareto distribution. For the proposed model it was assumed that distribution (density) of $h_s$ value is described with the Pareto distribution in form of following function.

$$f(x) = \frac{1}{x^{\alpha}} \tag{14}$$

Where:
$\alpha$ – coefficient of the distribution in the interval $[0,1]$

The number of passengers for who the $h_s < h_0$ is proportional to $P_1$.

$$P_1 = \int_{h_{min}}^{h_0} f(x)dx \tag{15}$$

The number of passengers for who the $h_s > h_0$ is proportional to $P_2$.

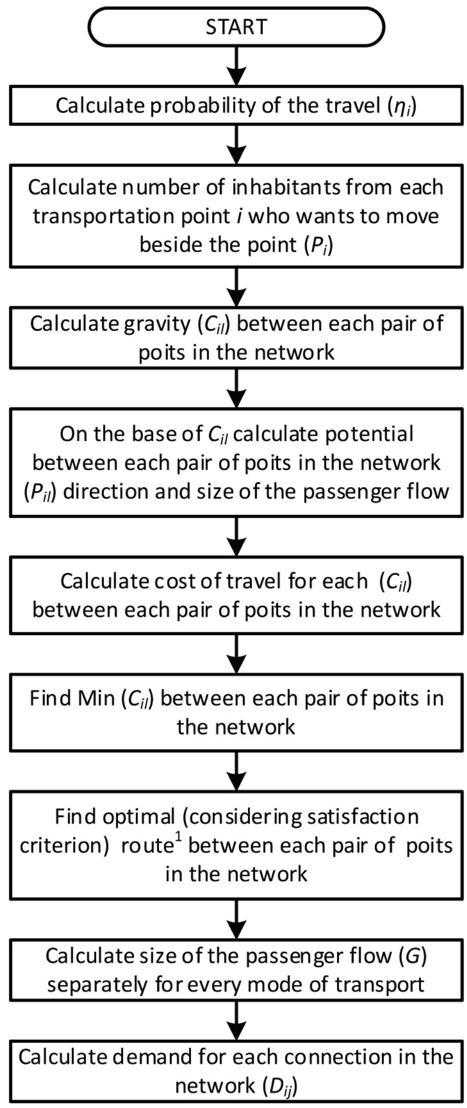
$$P_2 = \int_{h_0}^{h_{max}} f(x)dx \tag{16}$$

5. Calculating maximal passenger demand for each connection in the network ($D_{il}$). In this operation all the potentials resulting from the routs that use given connection $f_{il}$ are summarized.

$$D_{il} = \sum_{j=1}^{k} P_{jl} \tag{17}$$

Where:
$D_{il}$ - maximal demand for the connection between points $i$ and $l$.

The five stages of the data processing in the simulation tool that are further described in the mathematical model are shown in details in the block diagram in figure 10.




[1] it can be either public or private or combination of both

Fig. 10. Block diagram presenting implemented algorithm steps.

### *6.3. QUERIES, FORECASTING AND SCENARIO TESTING IN THE MODEL*

A set of scenarios were developed to elaborate how does the model behave in different circumstances. There is some application of the developed model. It can show results of any modifications in the transportation network regarding new transportation mode, change in the schedule, technology change influencing shuttles speed. The model can also be used as a forecasting tool which predicts passengers flow between chosen transportation points in the future up to the year 2025. There are three scenarios available for forecasting – optimistic, neutral and pessimistic. The model also provides optimal connection according to defined boundary conditions. The model's user can choose a train connection, a bus connection, a car connection, or a mixed bus-train connection. Subsequently, the user selects whether to optimise the distance, price, time or the relation price to time. Finally, the user can search for all routes, route from a selected point or a route from one to another transportation point. Figure 11 presents the exemplary effect of a search in the case when the user specifies two particular transportation points. An example shows a situation in which the user of a model wants to find a route (e.g. mixed bus-train) from Wroclaw to Jelenia Gora so that the connection is optimal when it comes to the price to time relation. The system suggests train connection on the complete route.


Fig. 11. Result of route search between two specified transportation points.

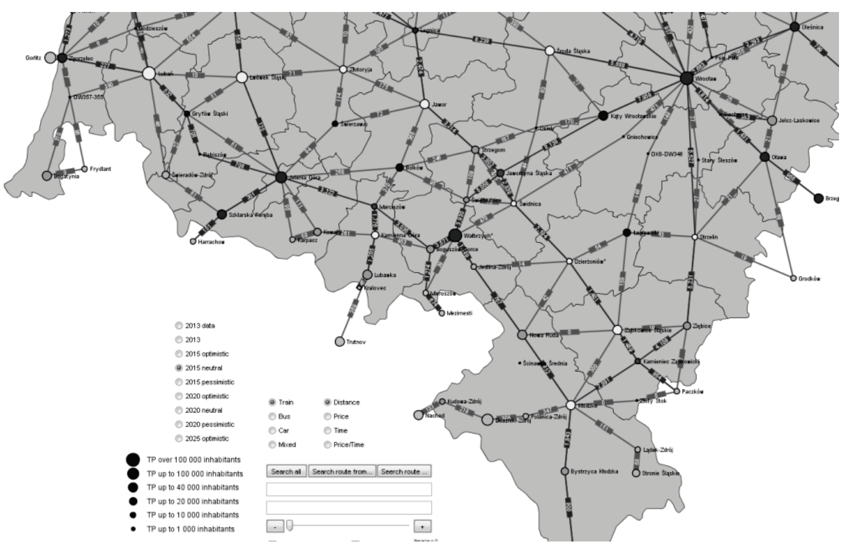



Fig. 12. Passenger flow for buses and trains – forecast for 2015 (neutral scenario).

Figure 12 shows a result in the case of a query about the forecast for a future, which is another possibility of the model application. It is the neutral forecast of the demand for train transport (red lines) and bus transport (grey lines) for 2015.

The proposed tool was designed as an advisory system for authorities responsible for the transport in a given region. There were presented two maps below to prove its usefulness in this field. They show preferences of passengers in an area of

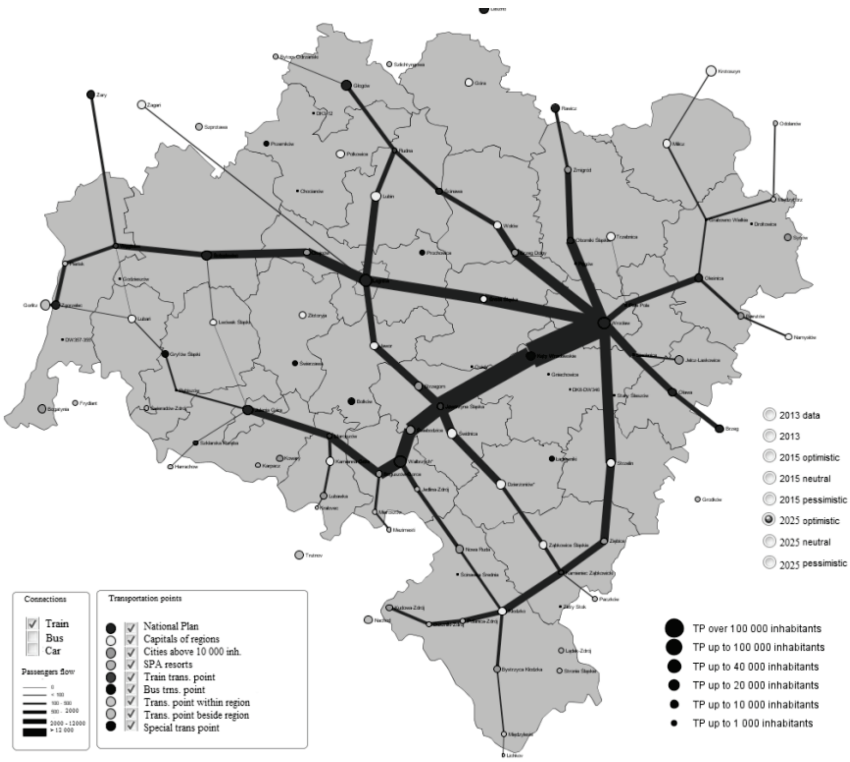



Fig. 13. Passenger flow forecast for railway (2025 – optimistic).

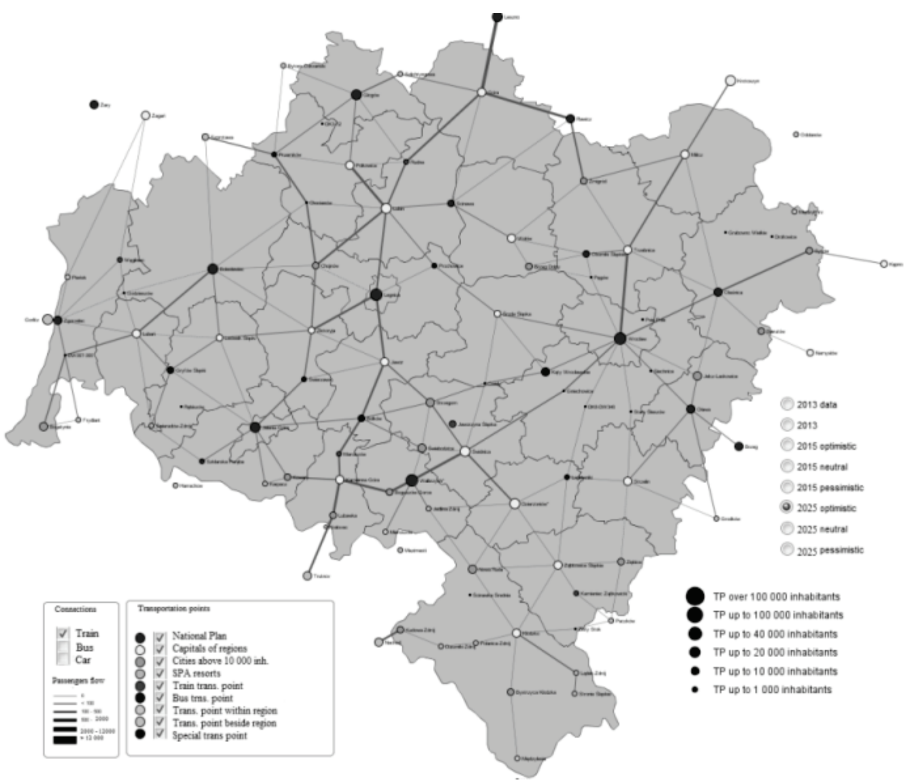



Fig. 14. Passenger flow forecast for buses (2025 – optimistic).

transport mode selection. Figure 13 demonstrates the map of passenger potentials forecasted for the year 2015 for railway (optimistic scenario). The thickness of the line indicates the size of the passenger potential. It is explained in details on the legend.

Figure 14 presents forecasted passengers flow for buses in the year 2025. It is also optimistic scenario.

The results of the analysis conducted in the simulation tool on the basis of real data enabled to verify its effectiveness and reliability. The generated forecast was evaluated as a valuable source of information that will be used to the transportation network planning and development in the region. The flexibility and universality of the model enable to use it as an advising tool for any configuration of transportation points with different means of transport. The conditions and data processing procedures can also be easily modified to some extent.

## 7. SUMMARY

The proposed tool was designed as an advisory system for authorities responsible for the transport in a given region. It gives the opportunity to plan a proper number of the public transport means and predict a load of roads considering car traffic. The results of querying and forecasting scenarios are clearly displayed on the maps in the simulation tool. The results of the analysis conducted on the basis of real data enabled verification of its effectiveness and reliability. The generated forecast was determined to be a valuable source of information that will be used to assist transportation network planning and development in the region. The flexibility, universality and adaptability of the model enables its use as an advising tool for any configuration of transportation points with different means of transport concerning any region or broader space. The conditions and data processing procedures can also be easily modified to some extent.

## REFERENCES


[1] R. Balcombe (edit.), *The demand for public transport: a practical guide*, TRL Report TRL593, 2004.

[2] D. Besanko, R. Braeutigam, *Microeconomics 4th ed.*, Wiley, 2010.

[3] A. Daly, *Estimating Choice Models Containing Attraction Variables*, Transportation Research, Part B: Methodological, Vol. 16, No. 1 (1982) pp. 5-15.

[4] R. M. Deshabrata, K. R. Sankar, P. B. Mahendra, *Multi-choice stochastic transportation problem involving extreme value distribution*, Applied Mathematical Modelling, Volume 37, Issue 4 (2013) pp. 2230-2240.

[5] J. Fessard, *Transportation Generalized Cost Functions for Railroads and Inland Waterways*. Massachusetts Institute of Technology, 1979.

[6] B. Jourquin, L. Tavasszy, L. Duan, *On the generalized cost - demand elasticity of intermodal container transport*. European Journal of Transport and Infrastructure Research, Issue 14(4) (2014) pp. 362-374.

[7] T. Nagatani, *Nonlinear-map model for bus schedule in capacity-controlled transportation*, Applied Mathematical Modelling, Volume 37, Issue 4, 2013, pp. 1823-1835.

[8] C. V. Phani Kumar, D. Basu, B. Maitra, *Modeling Generalized Cost of Travel for Rural Bus Users: A Case Study*. Journal of Public Transportation, Vol. 7(No. 2) (2004) pp.59-67.

[9] L. Tavasszy, I. Davydenko, K. Ruijgrok, *The Extended Generalized Cost Concept and its application in Freight Transport and General Equilibrium Modeling*, in: s.n. (ed.) Integration of Spatial Computable General Equilibrium and Transport Modelling, Tokyo, 2009.

[10] M. West, *Statistical Inference for Gravityels in Transportation Flow Forecasting*, Technical Report Number 60: National Institute of Statistical Sciences, 1997.

[11] A. Żurkowski, *Modelowanie przewozów międzyaglomeracyjnych (eng. Interagglomeration Transport Modelling)*, Problemy kolejnictwa (eng. Railway Issues), 53(148) (2009) pp. 5-47.


Date submitted: 2016-07-01
Date accepted for publishing: 2016-12-05

**Joanna Gąbka**
**Wroclaw University of Technology, Poland**
**joanna.gabka@pwr.edu.pl**

**Sławomir Susz**
**Wroclaw University of Technology, Poland**
**slawomir.susz@pwr.edu.pl**

**Maria Rosienkiewicz**
**Wroclaw University of Technology, Poland**
**maria.rosienkiewicz@pwr.edu.pl**