國立交通大學

交通運輸研究所

碩士論文

新興國家高速鐵路路線：以烏克蘭為例

Planning a High Speed Rail Route in an Emerging Country: A Case Study of Ukraine

指導教授：　馮正民老師

林楨家老師

研 究 生：　韓安東

中 華 民 國 一0二 年 六月

新興國家高速鐵路路線：以烏克蘭為例

Planning a High Speed Rail Route in an Emerging Country: A Case Study of Ukraine

指導教授：　凴正民老師
　　　　　　林楨家老師
研 究 生：　韓安東

Advisors: Professor Cheng-Min Feng
Professor Jen-Jia Lin
Student: Anton Hagen

國立交通大學
交通運輸研究所
碩士論文

A thesis
Submitted to Institute of Traffic and Transportation Management
College of Management
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in

Traffic and Transportation

June 2013

Taipei, Taiwan, Republic of China
中華民國一０二年六月

Chinese Abstract

烏克蘭為歐洲大陸中幅員廣大的國家之一，具有主要城市彼此座落甚遠的特性。目前其陳舊之鐵路網絡，僅能提供耗時的日間服務或隔夜班車，從首都基輔至西部、東部和南部的市中心需要八至十二小時。這不僅限制了國境內交通機動性，且成為區域發展的障礙。雖然烏克蘭鐵路在 2012 年歐洲足球錦標賽前進行改善，但其引進的 IC +列車並未解決烏克蘭鐵路所面臨的挑戰。由於引進高速鐵路(high speed rail, HSR)是已開發國家為運輸改善政策主流，故本研究著重建立針對烏克蘭高速鐵路的路線規劃方法。其研究目的為以下兩點：1.引發烏克蘭國內對高速鐵路規劃的科學性方法的討論，並進一步提出適於烏克蘭高速鐵路計劃的方法；2.本研究將規劃連接基輔與烏克蘭最東部的商業中心－頓涅茨克之高速鐵路路線。

本研究的規劃方法分為兩個階段：方案產生(alternative generation)和方案評估(alternative evaluation)。第一階段為列出鐵路路線的替選方案，此階段目的為列舉可能的路線及定義排除不可行方案的準則。本研究透過電腦模擬列出所有可能的路線，並採用階層分析法（analytic hierarchy process , AHP）及理想解類似度順序偏好法（technique for order of preference by similarity to ideal solution , TOPSIS），以選擇替代路線做進一步的評估。第二階段使用運輸需求分析（transportation demand analysis , TDA）和地 理資訊系統（geographic information system , GIS），針對每條路線的特徵進行分析，最終確定最適路線。

本研究的主要貢獻為以下兩點：1.提供一高速鐵路路線規劃之方法；2.並依據建議參考之評選準則提出最佳之高速鐵路路線規劃。所建議之最佳路線為連結基輔、切爾卡瑟、克里沃羅格、聶伯城及頓涅茨克等城市。本研究所提出之方法與成果可作為後續研究之參考，後續研究者可依據本研究之基礎提出進一步之觀點及成果，或是發展相關議題，如本研究建議路線之可行財務方案。

關鍵詞：高速鐵路、路線規劃、多準則決策。

English Abstract

Ukraine is one of the largest countries of Europe and its major cities are situated far away from each other. The current railway network is obsolete and provides either overnight or time-consuming daytime service, so it takes 8-12 hours to get from the capital city of Kyiv to the regional centers in the West, East and South. This limits mobility inside the country and creates obstacles for regional development. Although Ukrainian railroads were improved before the UEFA Euro 2012 Championship, introduction of IC+ trains did not solve challenges of Ukrainian railroad. As *high speed rail* (HSR) introduction is a mainstream in transportation improvement policies in developed countries, this research makes an attempt on developing a route planning approach for Ukrainian HSR. This study aims at two objectives: to start a scientific discussion of HSR in Ukraine, and to propose a method, that can be used in further HSR planning in Ukraine. The HSR line offered will connect Kyiv with Donetsk, the most eastern business center of Ukraine.

This study develops the planning approach in two main phases: alternative generation and alternative evaluation. The first phase generates route alternatives. The most important components of this phase are identifying possible routes and defining rules to cut off infeasible alternatives. The possible routes are identified via a computer simulation that generates the full list of the routes available. Using *analytic hierarchy process* (AHP) and *technique for order of preference by similarity to ideal solution* (TOPSIS), the alternative routes are selected for further evaluations. The second phase uses *transportation demand analysis* (TDA) and *geographic information system* (GIS) to analyze the characteristics of each route and finally determine the optimal route.

This study results in two contributions: it introduces a method of HSR route planning and recommends the optimal HSR route considering the chosen criteria. The route connects Kyiv, Cherkasy, Kirovohrad, Kryvy Rih, Dniprodzerzhyns'k, Dnipropetrovs'k, Zaporizhzhya and Donetsk. The proposed methods and results of this study can be used for the further studies in this field. Future researchers can either improve the ideas and results of this study or develop neighboring issues such as possible financing scenarios according to the results of this study.

Keywords: High speed rail, route planning, multiple criteria decision making.

Acknowledgement

I would like to express my deep gratitude to my advisors Professor Cheng-Min Feng and Professor Jen-Jia Lin for their patient guidance, enthusiastic encouragement and useful critiques of this study. I would also like to thank the experts who spent their valuable time on filling in the AHP surveys.

I would like to extend my thanks to my classmates for introducing me the life in Taiwan helping me to overcome the language barrier.

Finally, I wish to thank my parents and my sister for their support and encouragement throughout my study.

Contents

List of Tables

List of Figures

I. Introduction

1.1 Motivations and background

Ukraine is an emerging country of Eastern Europe, former USSR. According to CIA World Factbook, it is the second largest country of Europe (and the first one, if Overseas France is not considered); its square is 603.6 thousand km^2, east-west distance – 1300 km, and north-south distance – 900 km. The population of Ukraine is around 46 million of people. As this research will often consider socio-economic geography of Ukraine, it is important to provide here more information about the country, as shown in Figure 1.

Background map source: http://worldmap.org.ua/
Figure 1 Approximate socio-economic pattern of Ukraine.

Socio-economic pattern of Ukraine varies a lot from East to West and from North to South. It is caused by heavy industries that are concentrated in the south-east of Ukraine. Those industries where formed during 1930-s and stimulated population increase in the existing cities and created new ones. So, along with high industrialization, east of Ukraine can be characterized as highly-urbanized. There are also smaller places with industry in the center and west of Ukraine (vehicle production, chemical industries), that explains high urbanization level in Lviv region, for example, but most of chemical industries are now closed. The north-east and center are mainly agricultural areas and the north-west is covered with forests, so the population density is low there.

Table 1 Top-10 cities, by population (1 Jan. 2012)

Cities	Population
Kyiv (capital)	2 814 258
Kharkiv	1 441 362
Odesa	1 008 162
Dnipropetrovs'k	999 577
Donets'k	955 041
Zaporizhzhya	772 627
Lviv	729 842
Kryvy Rig	660 203
Mykolaiv	497 032
Mariupol	464 457

Although major cities are mainly situated in the east and center of Ukraine, there are still quite large and economically important cities in the west. This leads to the problem of intercity transportation in Ukraine: large cities are situated far away from each other. For instance, Lviv and Donetsk (western and eastern regional centers respectively) are 1.200 km away from each other.

At the moment, the main way of intercity transportation in Ukraine is an obsolete railway network, originally built for freight transportation. This causes the following weaknesses of transportation system:

1. Average speed is quite low, the fastest daytime train has average speed of 116 km/h, the fastest overnight trains – around 60-65 km/h. This leads to very long travel times (for case of Lviv-Donetsk travel, it will be 18-20 hours).
2. Railroad routes are not direct and/or deviate from large cities.

Because of low speeds and long distances, overnight travelling is widespread in Ukraine. Still, though this model is relatively comfortable for 8-10 hours trip, it is almost unacceptable for 20 hours trip. These weaknesses break social and economic relations inside the country and strongly limit business activities.

In order to improve railway services, the government developed a program of InterCity+ (IC+) traffic on existing tracks, but there was no public discussion; therefore, no information if the government considered HSR alternative for Ukraine is available.

During preparation to Euro-2012 championship, the major directions of Ukrainian railroad system were improved to allow continuous maximum speed of 160 km/h and new rolling stock from Hyundai Rotem was purchased (10 trains). At the moment these trains are commuting on routes Kyiv-Kharkiv (3 per day), Lviv-Kyiv, and Kyiv-Donetsk (2 per day).

Nevertheless, these changes haven't improved the situation much: only services of Kyiv-Kharkiv are relatively successful and, for other two routes, the following shortcomings exist.

1. Number of stops

Existence of intermediate stops is often very important: it helps to generate demand for transportation. Roughly, if you have route A to B with a stop C, there will be two types of traffic of each section (AC: A -> C, A -> B, CB: C -> B; A -> B). Currently Lviv route doesn't have intermediate stops because there are no large cities along the used track.

2. Travel times

Travel time is one of the measures that drastically influences travel demand. If the train doesn't show significant difference from conventional rail, it will not be popular. For example, there is an alternative way from Kyiv to Lviv that has 3 more large cities, but it will increase travel time from 5 hours to 6.5 hours. In the same time, the main trouble of Donetsk route is also travel time: it takes 6.5 hours to get to Donetsk.

3. Convenience

IC+ trains travel during daytime and it is not convenient for numerous people. For example, when business trip is necessary, employer will want his employee to return from a trip as soon as possible. If business requires a 2-day 9 a.m. - 18 p.m. presence in the other city, it means that the entire trip will take 4 days and 2 more nights in hotel. On the other hand, overnight train is much more convenient in term of time usage: the trip can start right in the end of working day and next morning that person is doing his job in another city.

4. Price

Ticket price for IC+ train is almost 90% higher than in premium-class overnight trains as shown in Tab. 2. In pair with relative inconvenience of IC+, this price scares away the passengers.

Table 2 Price comparison: IC+, Overnight premium and Overnight ordinary
Source: http://booking.uz.gov.ua

Route	IC+			Overnight premium			Overnight ordinary		
	Travel time	I class price	II class price	Travel time	I class price	II class price	Travel time	I class price	II class price
Kyiv - Lviv	4:55	39	26	7:52	22	11	9:42	14	9
Kyiv - Kharkiv	4:28	35	24	9:22	19	-	8:25	12	8
Kyiv - Donetsk	6:40	47	31	12:12	24	14	-	-	-

So, the introduction of Hyundai Rotem trains as IC+ did not fix the weaknesses of Ukrainian railroad system; therefore, this research remains needed for its improvement.

Of course, today's Ukraine needs a HSR network to connect Kyiv with the furthest regional centers and to connect them to each other, but every HSR project is very complicated in development and very expensive in building. So this study develops only one line. Considering current directions of IC+ trains, we can see that route Kyiv-Donetsk is very time-consuming, but very promising from the point of view of population covered.

This study offers construction of HSR line between two large Ukrainian cities: Kyiv (capital city) and Donetsk (business and industrial center of Eastern Ukraine). As Ukrainian cities are scattered throughout territory, there are different alternative routes possible. We will evaluate the best feasible alternatives, estimate their influence on regional development, and compare them with status-quo.

1.2 Research objectives

As it was already mentioned this research is developed to improve quality of passenger intercity transportation in Ukraine. The thesis has several purposes:

1. To define long-run development objectives for Ukrainian transportation system; it will show what should be emphasized during next few years and how can be improved railroad transport of Ukraine.
2. To create first scientific analysis for HSR in Ukraine. At the moment there is no information about any HSR project, so this research can give a start for scientific discussion about introducing HSR in Ukraine and related topics, such as HSR financing schemes.
3. To build and verify methodological framework that can be used for developing of another HSR routes in Ukraine, so that it will take much less efforts for the next researches.

1.3 Research scopes

1.3.1 Key terms

The key concept used in this thesis is **High Speed Railroad**, shortly **HSR**. There are different understandings of HSR, first of all because of different maximum speeds. In the study HSR is a railway system (infrastructure and rolling stock) that allows high speed movement up to 300 km/h and approximate average speed of 200 km/h.

On the contrary to HSR, **conventional railroad** is a basic railroad system, created during XIX-XX centuries with a priority in freight transportation. Consequently this railway network is bound to freight demand-generating nodes and is not fully efficient for passenger transport. Speed on conventional railroads is limited to 140 km/h, on major lines – to 160 km/h.

There are also two terms in this research, that sound very similarly - **HSR route** and **HSR line**, but they should not be mixed. **HSR route** is a route available for usage by high speed train at lower speed, while **HSR line** is an exclusive line for high speed trains. This division is necessary for the case of mixed HSR route, when it can include both HSR lines and conventional railroad lines.

1.3.2 Spatial and temporal scopes

This thesis aims at developing an approach to determine the optimal HSR line between Kyiv and Donetsk, which is the most eastern large business city. There are numerous possible routes, but the most probable are routes via Kharkiv (the largest city in the eastern Ukraine) or via Dnipropetrovsk and some smaller cities. Theoretically the most possible routes will lie in the highlighted area of Figure 2.

In terms of time, this project cannot be regarded as a short-run. In fact, HSR project is a financial and engineering challenge that means that every stage of its implementation will be time-consuming. For example, it took 7 years for Taiwan to build a 350-km HSR, so even if conditions in Ukraine are much easier, it will probably take about 15 years to build a 650-km line between Kyiv and Donetsk.

Still, the time of project implementation strongly depends on the type of financing that will be chosen for the project and from the route selected. For example, if the line will be built on the left side of the Dnieper River, it will require significantly less special structures like bridges and tunnels, because of flat terrain. It will reduce price and improve the speed of building. One more important example: in case line is compatible with conventional railroad (like in France or Germany), it will be possible to introduce HSR partially, temporarily including upgraded conventional rail into the system.

Background map source: http://worldmap.org.ua/

Figure 2 Area of possible routes from Kyiv to Donetsk.

1.4 Research process

As it was already mentioned, there are no alternatives available for evaluation right away and it strongly influences the structure of this study. It will involve evaluation techniques twice: to generate a list of 5 alternatives from the full list of routes and to evaluate the alternatives to obtain the best route.

This research is developed in 4 stages as shown in Figure 3.

<u>Stage 1</u> The first stage of this research aims to clarify the research purposes, scopes and works. Therefore, the data about Ukrainian geography, socio-economic pattern and current state of intercity transport is generalized on the first stage in order to create a general idea about the direction of the project.

<u>Stage 2</u> This stage is devoted to literature review in order to examine previous studies in the field of project. There are three directions of literature review:

- previous studies of the HSR projects in Ukraine;

- literatures that study HSR projects in the other countries;
- methods in the fields of transportation routing, demand estimation and project evaluation.

Stage 3 In this stage, the route planning method proposed in this study is described.

Stage 4 Case study

Stage 4.1 Alternative generation At first, a simplified routing algorithm is used that will produce all possible routes between the origin and the destination. Then they will be ranked by three criteria: route population, route length, and route curvature:

1. Route population (max). This criterion is an incentive for a HSR route to pass though maximum possible number of cities, because every new city on the route can generate new demand.
2. Route length (min). This criterion is an incentive for route length minimization because long route increases travel time and building cost.
3. Route curvature (min). This criterion is an incentive to make a HSR route as straight as possible, without covering the cities that are located far from the direct route between the origin and the destination.

As we are facing three-criteria optimization and those criteria have different preference direction, TOPSIS is used to balance these criteria and assign overall criteria satisfaction score to each alternative route.

Stage 4.2 Alternative evaluation Values for alternatives are evaluated.

Stage 4.2.1 Approximate transportation demand estimation for each feasible route using travel demand analysis.

Stage 4.2.2 Measurement of construction complexity criteria values.

Stage 4.2.3 Measurement of external effects criteria values.

Stage 4.2.4 Selection of the most optimal route based on the criteria of route travel demand, construction complexity, and external effects.

Stage 5 Conclusions and recommendations.

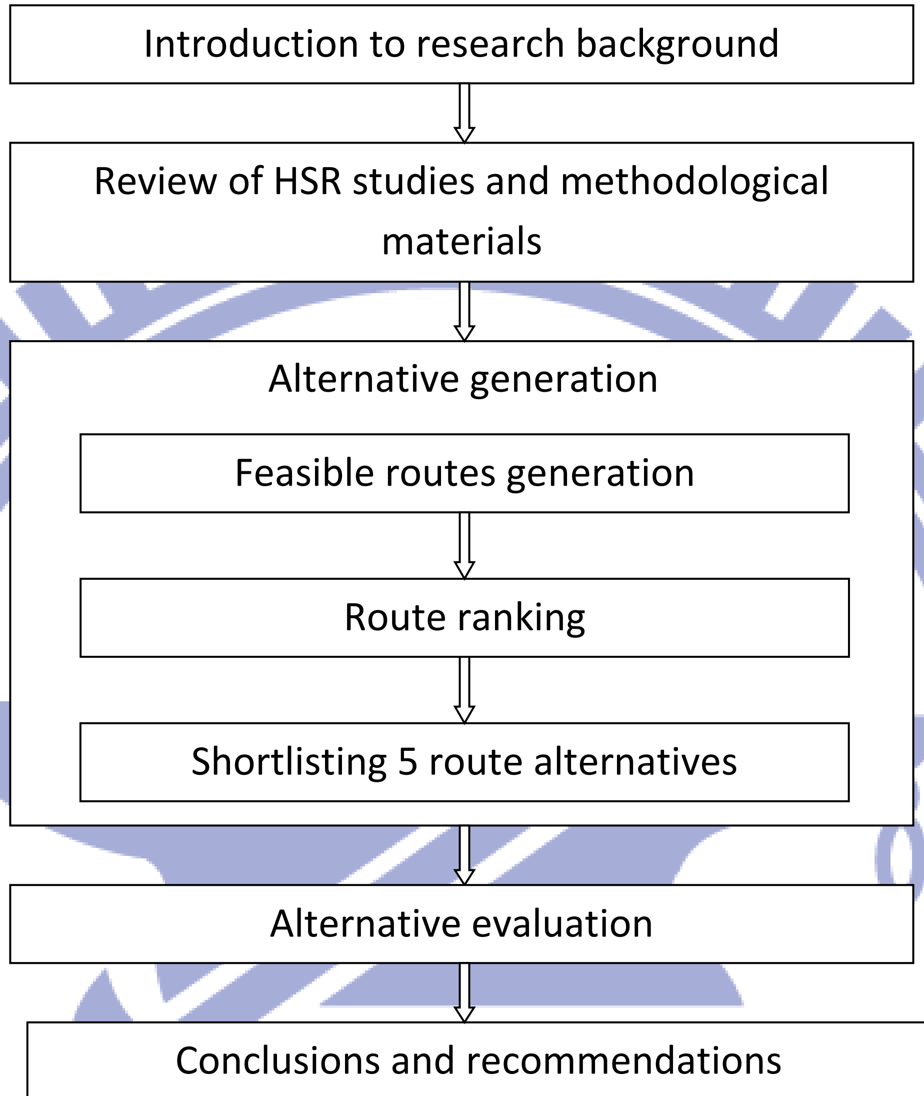


Figure 3 Research process

II. Literature Review

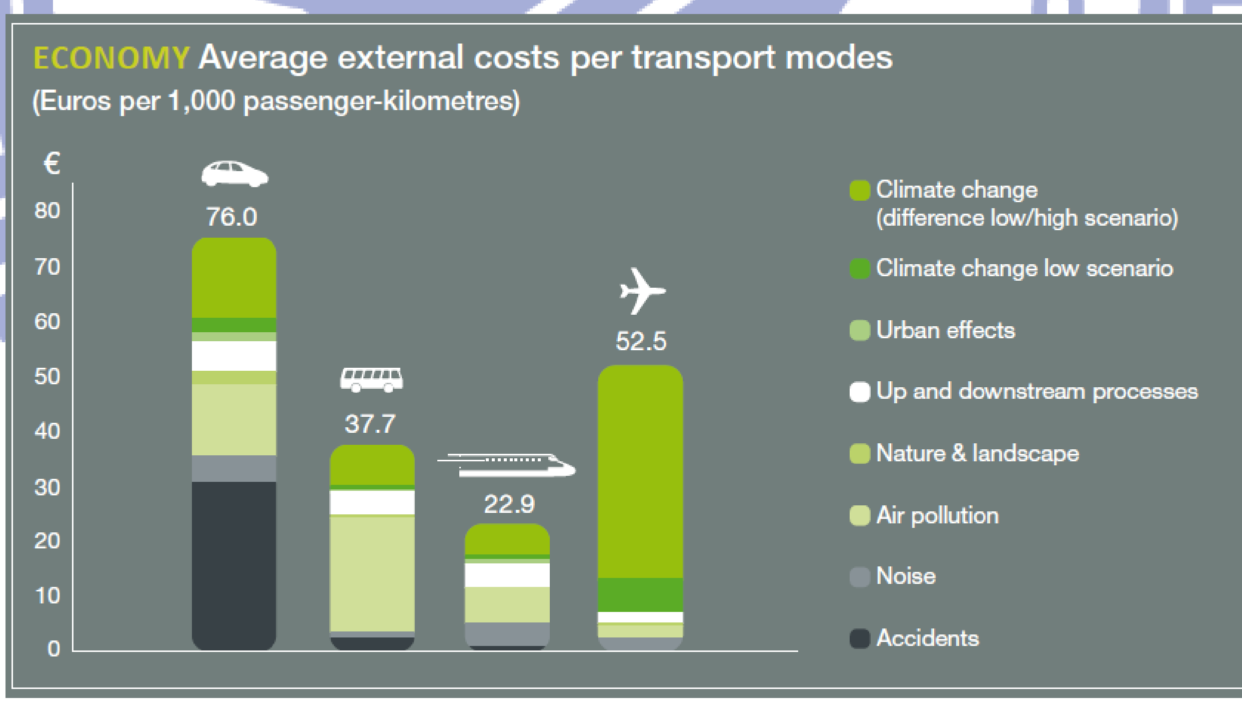
2.1 HSR studies review

Literature review of this research could include the following three topics:

- review of the literature, that directly concerns the topic of HSR development in Ukraine;
- review of studies, devoted to HSR development in different countries;
- general review of decision-making techniques about route planning.

Unfortunately, the first topic cannot be represented in this review: after investigations performed, no research papers were found on the topic. So, we assume that this research is the first known study of HSR in Ukraine and skip the first section of review.

Although there is no existing literature in Ukraine studied the topic, HSR development is a mainstream in contemporary transportation projects in developed countries. The phenomenon of railroad revitalization in the form of HSR was caused by carbon-reducing and energy-saving policy: for middle-length trips high speed electric-driven trains appeared to be more effective than aviation (Dobruszkes, 2011). In terms of CO_2 emission, emissions per passenger, produced by HSR are half of that travelling by car and quarter of emissions, produced by air transportation (Cascetta and Coppola, 2011).



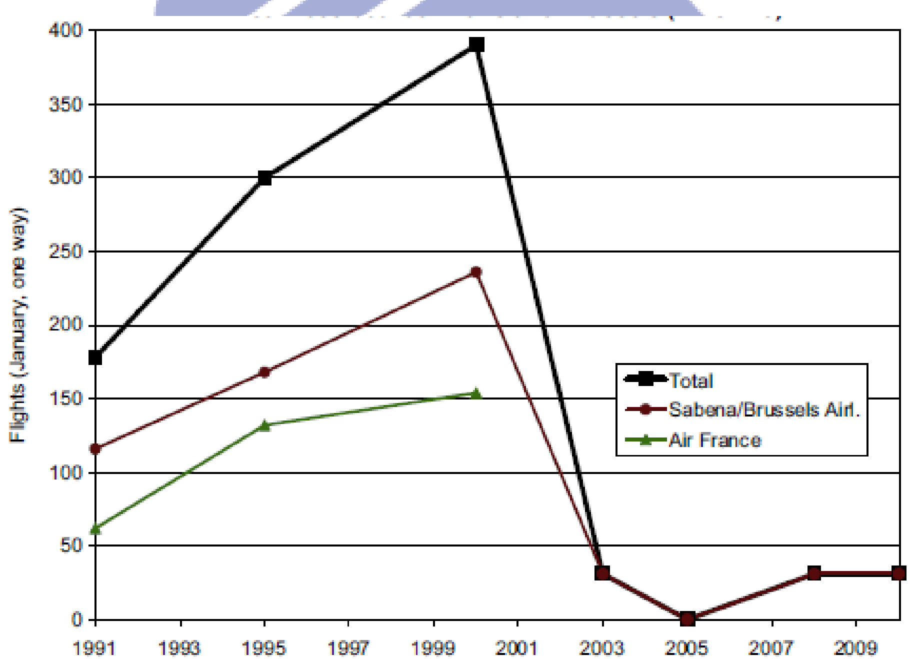
Source: UIC (2010)

Figure 4 Average external costs per transportation mode

As it can be seen on the Figure 4, there are strong incentives to encourage modal shift from aviation and cars to HSR lines, as it influences environment less that other modes.

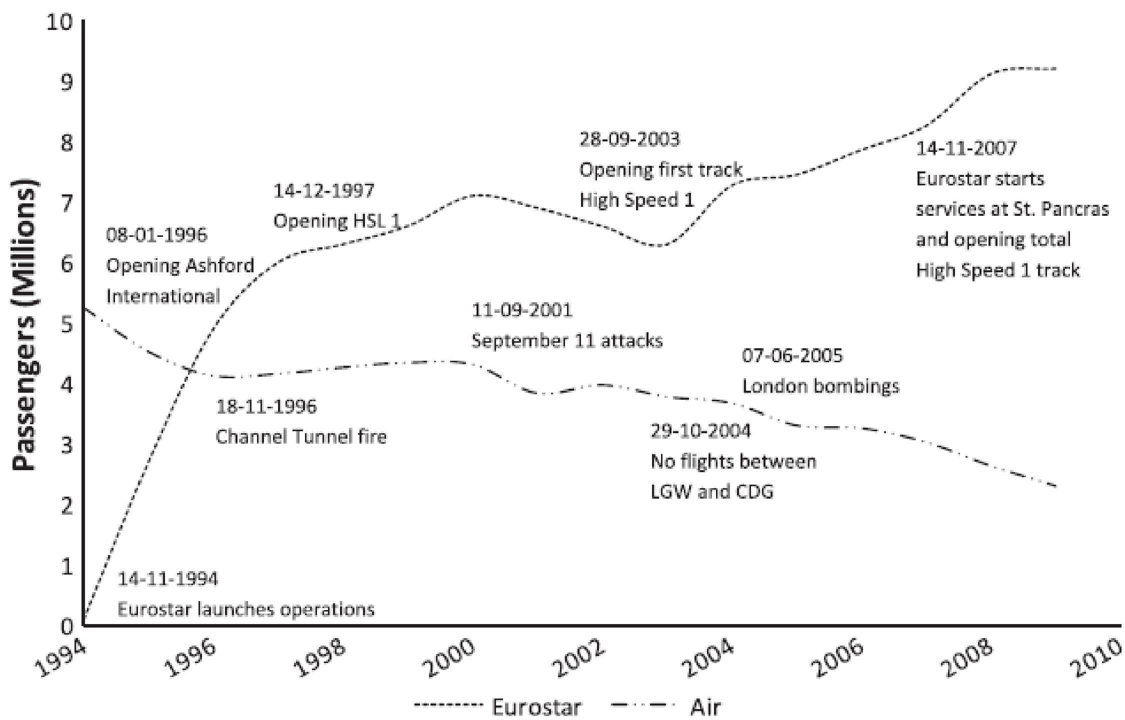
There is also a significant difference in the land use. While standard double-track railway line requires 25 m wide line, a 6-lane motorway requires 75 meters (the area occupied is 3.2 ha/km and 9.3 ha/km respectively), but the capacity is almost the same. Moreover, there is a common practice to build HSR lines parallel to existing motorway that helps to reduce land use significantly (UIC, 2010).

HSR benefits of high speed combined with high accessibility that both create substantive competition advantage in comparison to aviation: if the passenger wants to use plane, he/she should go to the airport that is situated outside the city and check-in the flight at least 45 minutes in advance before flight. If consider average travel time from airport to city as 30 minutes, "brutto" travel time will be 1 hour 45 minutes larger than net flight time. On the contrary, railway station can be reached in 15 minutes that makes HSR faster than aviation on the medium distances. Also, basing on UIC estimations, some modal shift assumptions can be made: for routes less than 800 km, HSR shifts 50% of passengers, for routes between 400 km and 600 km – 70%, and for routes shorter than 400 km modal shift will be 85-100% (Ehrenberger *et al.*, 2010). In fact, introduction of HSR can lead to serious changes in air services. For example, as shown on the Figure 5, after HSR service appeared on the links "Paris-Metz" and "Paris-Nancy" (both approx. 300 km) lead to complete discontinuation of air services on this link and the introduction of HSR between Paris and Brussels lead to elimination of all Air France flights and bankruptcy of Sabena Air Company, that did an attempt to compete with Thalys HSR (Dobruszkes, 2011).



Source: Dobruszkes (2011)
Figure 5 Air services between Paris and Brussels

Quite similar situation in on the link between London and Paris, but the aviation is not fully eliminated there, probably because of high fares for Eurostar trains and because London Heathrow airport often is not a final destination, but a hub (Dobruszkes, 2011). The competition between HSR and aviation on the link between London and Paris can be seen on the Figure 6.



Source: Behrens and Pels (2012)

Figure 6 Number of passengers for London-Paris link

HSR introduction causes not only modal shift on the OD pair, but also generates new demand, called induced demand. Changes in travel time and travel cost that can be referred as total trip cost change, they also can be a measure of accessibility (Yao and Morikawa, 2003). These changes in accessibility cause changes of socio-economic situation: changed travel times influence people's choice of place and work. For example, people can move to other working places or move themselves to more living-convenient rural areas, becoming commuters (Cascetta and Coppola, 2011). This is the reason, why HSR lines in many countries are regarded as a way of regional development and decentralization (Ryder, 2012). The amount of induced traffic is usually very significant, that causes strong interest towards this phenomenon. For example, in France and Japan HSR produced additional traffic as high as 35%, and the study in the Canberra-Sydney obtained a value of 26% even without taking in consideration changes in land use that can happen in the long-term (Hensher, 1997).

2.2 Methodology review

There is a significant lack of studies in the field of route planning for railroads. Usually scientists evaluate given alternatives or make some specific studies about the project that already exists, but there are no commonly used methods, for the case, when the routes should be generated first. Nevertheless, some common methods of transportation will be used.

Regarding transportation routes evaluation, the most widespread are the studies devoted to demand estimation, assessment of external effects, such as air pollution, noise and vibration pollution, and influence on living areas.

The most generally-used model for demand estimation is transportation demand analysis (TDA). It consists of 4 stages: trip generation, trip distribution, modal shift, and traffic assignment.

According to Caulfield (2011), **trip generation** refers to the amount of trips generated by each origin and destination that are influenced by several socio-economic feature of the zone. It includes two components:

- trip production (by the origin points), can be influenced by population, wages on the macro-level and by family size, social status, availability of car on the household level;
- trip attraction (by destinations) is usually assumed to be influenced by the type of land use, economic activity, employment.

Trip distribution, as a second step of TDA, may be formed in different ways, but the most widely used is gravity approach to the distribution of trips. It assumes that trips from each origin towards each destination are distributed according to Newton's gravity law as follows:

$$T_{ij} = \alpha O_i D_j f(c_{ij}) \qquad (1)$$

where T_{ij}: travel amount on the link between cities *i*, *j*;
α: proportionality factor;
O_i, D_j: trips, generated by origin *i* and destination *j*;
$f(c_{ij})$: generalized function of travel cost between *i* and *j*.

Generally speaking $f(c_{ij})$ here is a resistance factor, so not only the cost can be used at this place, but also travel time or distance.

Mode split is usually based on the utility of each alternative mode that consists of predicable value *V* and random value ε as follows:

$$u_m = v_m + \varepsilon_m \qquad (2)$$

where u_m: utility of mode *m*;

v_m: predicable utility component of mode *m*;
ε_m: random utility component of mode *m*.

Predictable component is a function of the transportation mode characteristics, such as travel cost, travel time, additional costs, etc. Finally, using the utilities calculated, mode choice model is built as a multinomial logit model as follows:

$$P_m = \frac{e^{v_m}}{\sum_i e^{v_i}} \tag{3}$$

where P_m: probability of choosing mode *m*;
v_m: predicable utility component of mode *m*.

The last step of the conventional model – **traffic assignment** – concerns link availability as a supply and number of O-D pairs and transportation modes as a demand. This step operates with a term of Level of Service (LOS) for each link of the network: the higher is the demand on each link, the worse are the traffic conditions and the larger is travel time. Mathematically:

$$t = t_0 e^{V/Q_S} \tag{4}$$

where t: travel time at the link;
t_0: travel time under free flow conditions;
V: flow;
Q_s: link capacity.

It should be taken into account that transferring an O-D pair from one link to another will simultaneously cause improvement of LOS on the former, and decline of LOS on the latter.

According to the purposes of analysis, this method can be used partially, depending on the detailing of analysis needed. Also, TDA can be used in several variations, depending on the data available. For example, Ehrenbreger et al. (2010) uses a TDA that joins stages 1 and 2 of the analysis for European HSR lines. This integrated model utilizes socio-economic data of two cities and passenger traffic on the link between them. The study develops a gravity model for European cities, where transportation demand can be estimated via GDP, tourism intensity, population size, and distance. In the final model tourism data is omitted, because it does not influence result.

$$F_{ij} = \beta_0 \cdot (P_i \cdot P_j)^{\beta_1} \cdot (W_i \cdot W_j)^{\beta_2} \cdot \left(\frac{d_{ij}}{\overline{v_{ij}}/100} \right)^{\beta_3} \tag{5}$$

where F_{ij}: travel amount between cities i and j;

$\beta_0 \ldots \beta_3$: coefficients;
P_i, P_j: population of cities i and j;
W_i, W_j: GDP of cities i and j;
d_{ij}: distance between cities i and j;
v_{ij}: average speed on the link between i and j.

We should note that although the above study is devoted to the development of second-generation HSR, it offers methods, useful for this research.

Evaluation of HSR route requires not only travel demand data, but also information about engineering and environmental feasibility of the project. Depending on the accuracy level needed and available data, methods can vary from rather precise monetary evaluation to rough assessment, but generally they usually use Geographic Information Systems (GIS).

Both kinds of feasibility refer to the construction complexity, and, consequently, to construction cost. According to Uršej and Kontić (2007), it is influenced by such factors:

- cost of special constructions, such as embankments, cuttings, bridges and tunnels;
- cost of living-areas protection, such as noise and visual barriers;
- cost of natural and cultural heritage protection: remediation of environment, creating crossings for animals;
- additional cost, such as related to land purchase.

The key idea of route assessment in the study of Uršej and Kontić (2007) is thesis, that HSR should go underground, only in the cases, when surface solution is unavailable due to surface space, causes great negative environmental impacts, or is more expensive than subsurface alternatives. That's why route planning starts form the suitability analysis of the surface where the alternative route should lie. Finally, alternative routes are ranked with respect to length of the route, number of tunnels and subsurface sections length.

GIS is also used in the study by Ehrenbreger et al. (2010) to evaluate technical complexity of the future line. This estimation is based on two ideas: there is a basic construction cost per kilometer and resistance multipliers that are calculated from resistance maps. At the resistance map, each raster pixel is assigned a resistance value depending on geographic conditions. Parameters that significantly influence construction cost are:

- Slope of the terrain.

As railway lines are restricted to the maximum gradient (for HSR maximum is usually 35-40‰ on the exclusive tracks, (UIC, 2010)), the cases exceeding this limit have to be handled by building of embankments or tunnels, bridges. Bridges and tunnels are assumed to be 7 times more expensive than the basic cost value.

- Areas with high population density.

A higher population density leads to the higher construction costs, because such areas have a few free areas for HSR line. So, the population density over 100 persons per pixel causes linear growth of the resistance multiplier. If the population is less than 100 persons per pixel, it does not affect construction cost.

- Water bodies.

Crossing of rivers is assign to have a resistance factor of 5.4, other water bodies cannot be crossed except the British channel. As it already has a tunnel, it has a small resistance factor of 1.5 (larger than 1 because of high fares).

More criteria for route evaluation can be found in California High Speed Rail Authority (2012) that develops HSR route in California, USA. One of the issues studied is a section of HSR between two cities and environmental impact of alternative routes. To evaluate the routes, two groups of criteria are introduced: physical and operational characteristics (such as travel time (minimization), intermodal connections (maximize), route length (minimize) etc.) and environmental impacts (air pollution, noise, vibrations, cultural and human hazards). Most of these criteria are already discussed, but there is one extra – intermodal connections, that helps to evaluate HSR line not only as a single transportation unit, but a part of transportation infrastructure of the region.

III. Route Planning Methods

This research uses different methods to obtain an optimal HSR route. These methods are: computer simulation, TOPSIS (technique for order of preference by similarity to ideal solution), AHP (analytic hierarchy process) and TDA (transportation demand analysis). The planning process is shown in the Figure 7 and explained as follows.

3.1 Computer simulation

Route evaluation studies usually concern an evaluation of existing route alternatives. In this study there are no given alternatives, so we have to generate them first. The literature studied does not give any method to generate alternative routes, so the route generation algorithm was developed specifically for this study.

To obtain route alternatives, we generate all possible routes between origin and destination; rank them according to the criteria and create a shortlist of 5 feasible rotes. Those feasible routes will be used further as alternatives for evaluation.

To generate full list of possible routes, this study uses recursive computer simulation to implement alternative generation algorithm that is based on the assumption that city population is a proxy to transportation demand it generates. Total sum of on-route city populations is used to roughly estimate total transportation demand of each route and rank the routes. Top routes will be considered as feasible and used for further analysis.

The classical TDA method is not used at this stage of the research, because its implementation for each route in the full list (that can theoretically contain hundreds and thousands of routes) will dramatically increase the computation complexity and time consumption of ranking, so that it will likely block the further progress of the study.

Simulation of the algorithm requires dataset with information about the cities. Each city record contains data about:

- City name
- Population

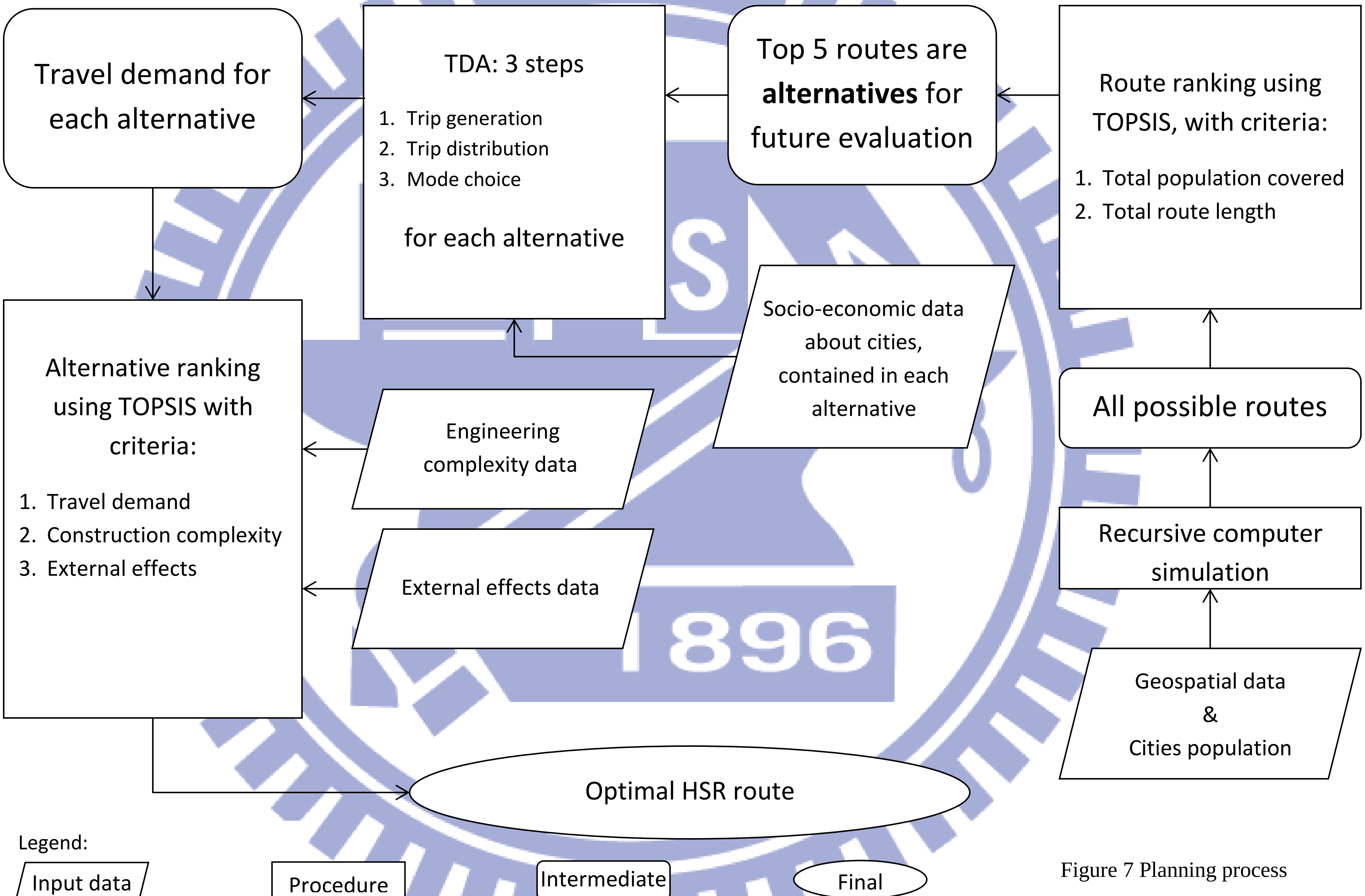


Figure 7 Planning process

In addition to this we need spatial information about the cities; this data is represented by distance matrix.

Let

- $C = \{0, 1, \ldots, i, \ldots n\}$: set of cities;
- $co, cd \in C$: are given origin and destination cities respectively;
- cc $\in C$: current city – a city where the recursion is located at each moment;
- $D = [d_{ij}]\ \forall\ i, j = 0..n, i \neq j$: a distance matrix, while d_{ij} is a distance between cities i and j;
- P_i: population of city i, where $i = 0..n$;
- T: population threshold (to exclude cities that too small to be considered in the simulation).

As the result of simulation we expect to have a set of routes $R = \{r_0, r_1 \ldots r_m\}$ where $r_i = (S, p, l, a)\ \forall\ i = 0..m$ is a route; p - total population, l - length and a - curvature index of each route. City list $S = (s_0, s_1, \ldots, s_n)$ is also a vector containing indexes of cities connected by the route i.

For running a recursion we need two additional variables associated with stack: recursive algorithm needs a “memory” to store current recursion depth and all previous steps. Let $stack = (S, pos)$ be a vector of city list and position variable, where city list is the same vector as for route $stack.S = (s_0, s_1, \ldots, s_n)$.

At the beginning we are situated at origin, so current city $cc = co$.

Recurrent function. If current city is not a route destination, then we add it to the current position of stack and launch the Search function (search rules will be defined further) of cities that can be next stop along a route. This search returns us a set of next possible node indexes NS = $\{i, i \in C\}$ For each element of the NS set we run this function recurrently again and again until we reach the stage, when “current city” will be destination point. In this case we record our stack to the next free route and return one step back in the stack.

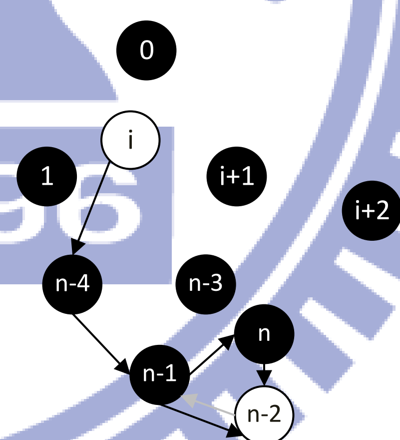


Figure 8 Search function step example.
Let we have to find routes between *i* and *n-2*. After 1st route is found, algorithm returns to the node, previous to destination and searches for another route

This algorithm is schematically described on the Figure 9; example of its intermediate result can be seen at Figure 9.

As it can be understood from Flowchart 1, recursion does not use given spatial and social data explicitly, it operates only with the next possible nodes, returned by the Search function.

Search function. A very important module for route generation is the Search function. It considers information about the origin city ($co \in$ C), current city ($cc \in$ C), and destination city ($cd \in$ C) and tries if any of cities $i \in$ C can be the next city on the route form *cc* to *cd*. The city is added to the output set of possible next cities NS, if all the following conditions are fulfilled:

1. $P_i \geq T$: we take into account only cities with population over threshold level;
2. $d[cc, cd] \geq d[i, cd]$: every next city should be closer to destination;
3. $d[cc, cd]^2 \geq d[cc, i]^2 + d[i, cd]^2$: obtuse angle between current city, next city and destination is required;
4. $d[stack.S_{pos-1}, i]^2 \geq d[stack.S_{pos-1}, cc]^2 + d[cc, i]^2$: obtuse angle between previous city, current city and next city is required;
5. $d[c_o, cd]^2 \geq d[co, i]^2 + d[i, cd]^2$: obtuse angle between origin, next city and destination is required.

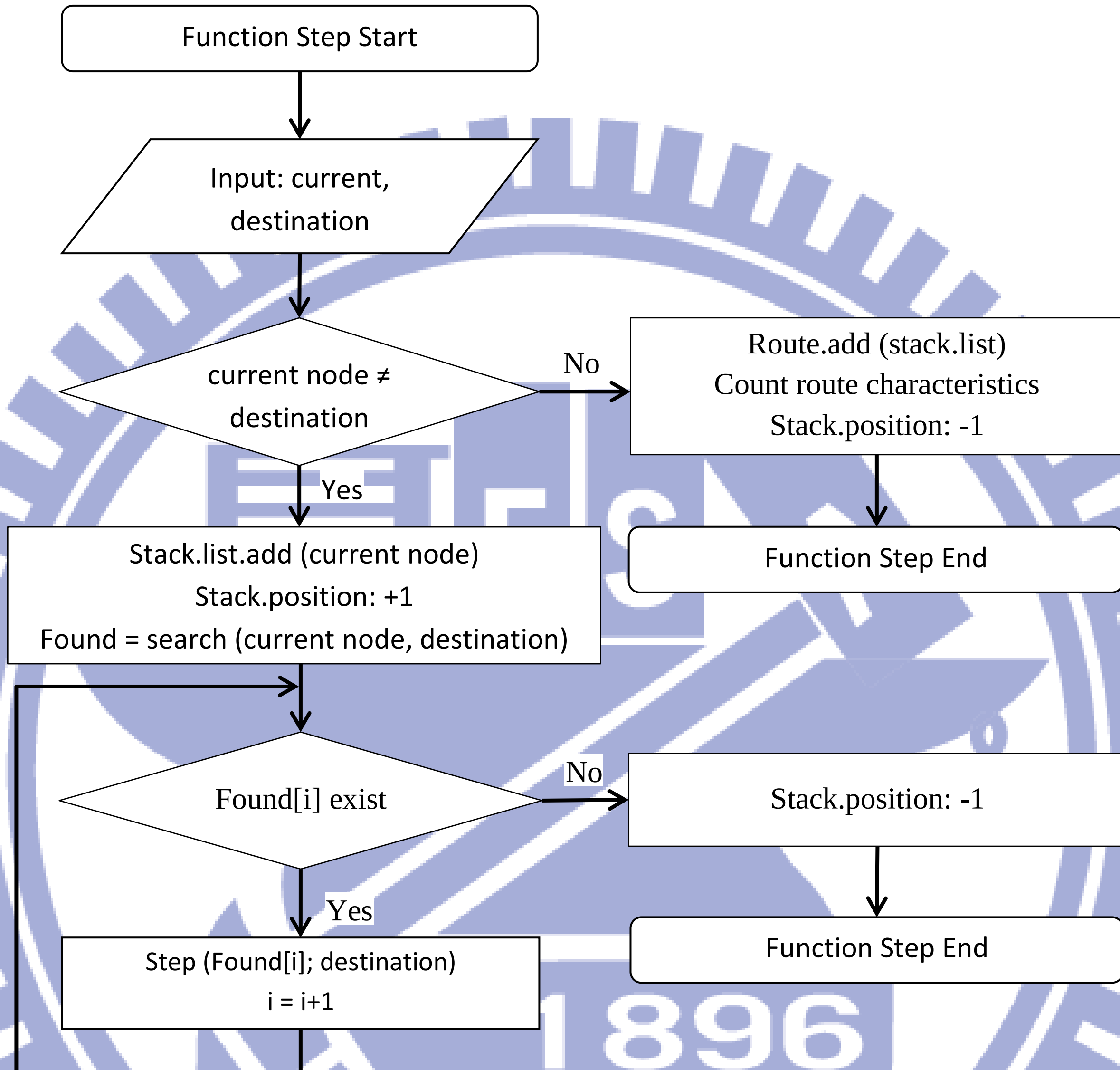


Figure 9 Route generation algorithm

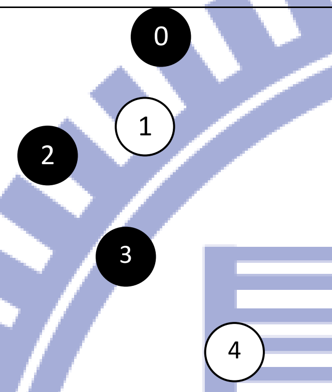
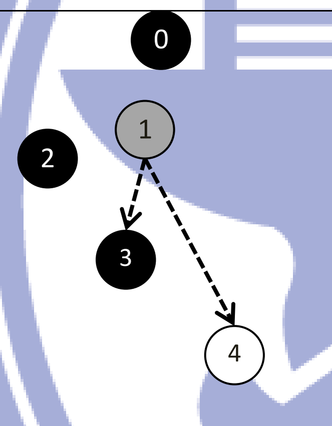
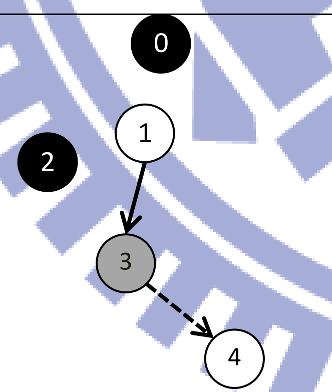
It can be quite difficult to understand, how the algorithm works, so it is useful to regard a small example as illustrated in the Table

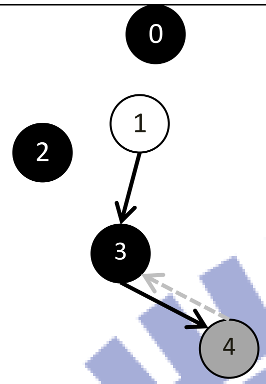
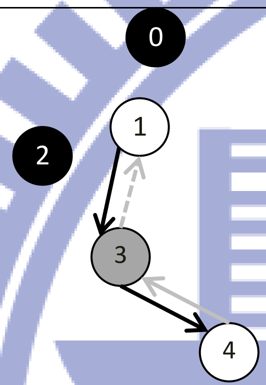
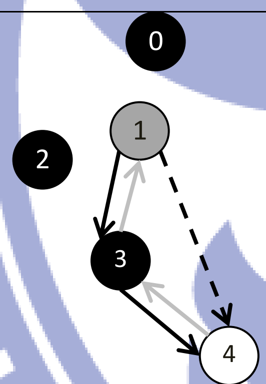
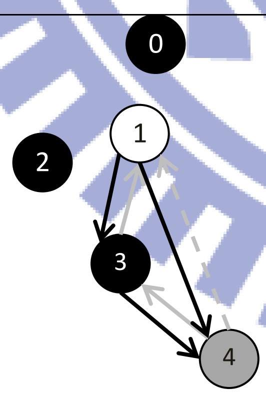
Having cities with indexes from 0 to 4, the algorithm finds all possible routes between given origin and destination subject to search rules. On the diagram you can see:

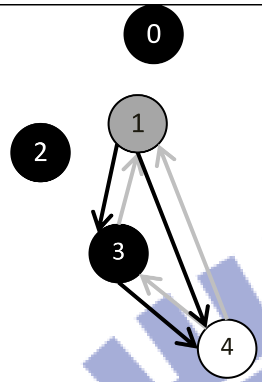
1. Circles with numbers as cities
    a. Black – ordinary city
    b. White – terminal city
    c. Grey – current city at a given step

2. Arrows witch are interpreted as connections between the cities:
    a. Black dashed – possible forward-steps at current city
    b. Black solid – previous forward-steps
    c. Grey dashed – backward-steps at current city
    d. Grey solid - previous backward-steps at current city

Table 3 Route search example

0		Let: Origin = 1 Destination = 4 Calling Function Step (1, 4) **Algorithm started.**	Stack.list = {1}; Stack.position = 0;
1		Function Step (1, 4) First, algorithm searches all possible next points of the route. According to the search rules, cities 0 and 2 are not qualified to the next step (city 0 because of rule 2, city 2 because of rule 3) , so the set NS = {3, 4} Next origin is city 3 (NS_0), call function Step (3, 4)	Stack.list = {1, 3}; Stack.position = 1; Routes = {}; current_city = 1;
2		Function Step (3, 4) According to the search, only city 4 is qualified to the next step, NS = {4} Next origin is city 4, call function Step (4, 4)	Stack.list = {1, 3, 4}; Stack.position = 2; Routes = {}; current_city = 3;

3		Function Step (4, 4) As here origin = destination, we record current stack as a route and count one position back. As this function was called in (2), we return to Function Step (3, 4)	Routes = {{1, 3, 4}}; Stack.position = 1; Stack.list = {1, 3}; current_city = 4;
4		Function Step (3, 4) NS = {4} NS_0 was already tested, so we step to NS_1. It is empty, so we go back one stack position more. As this function was called in (1), we return to Function Step (1, 4)	Stack.position = 0; Stack.list = {1}; current_city = 3;
5		Function Step (1, 4) Having NS = {3, 4} NS_0 was already tested, so we step to NS_1. Next origin is city 4, call function Step (4, 4)	Stack.list = {1, 4}; Routes = {{1, 3, 4}}; Stack.position = 1; current_city = 1;
6		Function Step (4, 4) As here origin = destination, we record current stack as a route and count one position back. As this function was called in (1), we return to Function Step (1, 4)	Routes = {{1, 3, 4}, {1, 4}}; Stack.position = 0; Stack.list = {1}; current_city = 4;

7	 	Function Step (1, 4) Having NS = {3, 4} NS_0, NS_1 were already tested, so we step to NS_2. It's empty, we exit function Step (1, 4). **Algorithm terminated**.	Routes = {{1, 3, 4}, {1, 4}}; Stack.position = 0; Stack.list = {1}; current_city = 1;

So, for this example the algorithm produced 2 routes – route 1 including cities {1, 3, 4} and route 2 including cities {1, 4}.

This algorithm will return a list of all possible routes between origin and destination, where each route includes city list and route characteristics used as criteria for ranking.

3.2 TOPSIS

In this study multiple criteria evaluation method TOPSIS will be used twice:

- Level 1 evaluation. After full list of routes is generated, it should be ranked according to the chosen criteria to select 5 feasible routes.
- Level 2 evaluation. After transportation demand and other criteria for final evaluation are calculated, TOPSIS is used to select the best of 5 feasible routes.

TOPSIS is extensively used because it gives a mechanism to rank alternatives according to multiple criteria of different nature with different preference direction (either maximization, or minimization). Available alternatives with the values of criteria are written as a decision matrix, see Table 4

Table 4 Decision matrix

e_{ij}		Preference direction	j				W_i
			A_1	A_2	…	A_n	
i	C_1	{MIN|MAX}	e_{11}	e_{12}	…	e_{1n}	w_1
	C_2	{MIN|MAX}	e_{21}	e_{22}	…	e_{3n}	w_2
	…	…	…	…	…	…	…
	C_m	{MIN|MAX}	e_{m1}	e_{m2}	…	e_{mn}	w_m

where $A_1 \ldots A_n$ are alternatives;
$C_1 \ldots C_m$ are criteria;
e_{ij} are values of criteria C_i of the alternative A_j.

$w_1 \ldots w_m$ are weights of the criteria.

As initially values of criteria are not comparable, so the matrix should be normalized:

$$r_{ij} = \frac{e_{ij} - e_{i\min}}{e_{i\max} - e_{i\min}} \quad (6)$$

where e_{ij} : values of criteria C_i of the alternative A_j.
e_{imax}: maximum value among the alternatives A_j for the criteria C_i.
e_{imin}: minimum value among the alternatives A_j for the criteria C_i;
r_{ij}: normalized values of criteria C_i of the alternative A_j. .

Then each r_{ij} is multiplied corresponding weight w_i

$$v_{ij} = w_i r_{ij} \quad (7)$$

where r_{ij}: normalized values of criteria C_i of the alternative A_j;
v_{ij}: updated values of criteria C_i of the alternative A_j;
w_i: weight of the criteria C_i.

According to preference direction, minimum and maximum values of each criterion are taken to create vectors of positive ideal and negative ideal solutions. For example, if there are 3 criteria with preference direction of MAX, MIN, and MAX, the positive ideal solution will contain the maximum value of C_1, the minimum value of C_2, and the maximum value of C_3. Two vectors obtained to compute Euclidian distance between every alternative and both positive ideal and negative ideal solution.

$$A_{j+} = \sqrt{\sum_i (p_i - v_{ij})^2} \quad (8)$$

$$A_{j+} = \sqrt{\sum_i (n_i - v_{ij})^2} \quad (9)$$

where n_{ij}: negative ideal solution for criteria C_i;
p_{ij}: positive ideal solution for criteria C_i;
v_{ij}: updated values of criteria C_i of the alternative A_j.

Finally, the score of each alternative A_j is computed as:

$$S_j = \frac{A_{j-}}{A_{j-} + A_{j+}} \quad (10)$$

For the purpose of making route alternative generation more automated, route ranking by TOPSIS method (Level 1) in this study is included into the alternative generation software; user only has to input weights of criteria.

3.3 AHP

TOPSIS method, previously mentioned, requires weights of criteria as input. To obtain these values, AHP is used. It includes several steps:

1. AHP survey. Experts are asked to perform pair wise comparison of the criteria.

For 3 criteria questionnaire field will look like on the Table 5.

Table 5 Pair wise comparison table

	9:1	8:1	7:1	6:1	5:1	4:1	3:1	2:1	1:1	1:2	1:3	1:4	1:5	1:6	1:7	1:8	1:9	
C1																		*C2*
C1																		*C3*
C2																		*C3*

2. Next, the comparison matrix is configured as shown in the Table.

Table 6 Comparison matrix

	C_1	C_2	C_3
C_1	1	c_{12}	c_{13}
C_2	$1/c_{12}$	1	C_{23}
C_3	$1/c_{13}$	$1/c_{23}$	1

3. The comparison matrix is normalized by dividing each element by the column sum.
4. Averaging the rows provides the weights of the criteria.

After obtaining the result it is useful to check the consistency of judgments made by the experts. To do this, next steps are followed:

1. Each column of the original comparison matrix is multiplied corresponding weight as follows:

$$\begin{pmatrix} 1 \\ 1/c_{12} \\ 1/c_{13} \end{pmatrix} (w_1) + \begin{pmatrix} c_{12} \\ 1 \\ 1/c_{23} \end{pmatrix} (w_2) + \begin{pmatrix} c_{13} \\ c_{23} \\ 1 \end{pmatrix} (w_3) = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \tag{11}$$

2. Then for each criterion consistency is calculated.

$$\forall i : cc_i = \frac{a_i}{w_i} \tag{12}$$

3. Calculation of consistency index CI:

$$CI = \frac{\sum_{i=1}^{n} cc_i / n - n}{n-1} \tag{13}$$

where n is the number of criteria.

4. Finally, consistency ratio is calculated and compared to the maximum value of 0.1:

$$CR = \frac{CI}{RI} < 0.1 \Rightarrow consistent \qquad (14)$$

where RI is taken from the Random Index table according to the number of criteria.

3.4 Transportation demand analysis

Next important method that will be used is transportation demand analysis. In general, transportation demand analysis consists of 4 steps:

1. Trip generation. Provides trips, generated by population, employment, incomes, centers of attractiveness in origins and destinations.
2. Trip distribution. Estimates the demand for transportation for each OD pair by using gravity model for traffic generated by O, traffic attracted by D and travel time between O and D.
3. Modal split. Evaluates modal shares for each OD pair.
4. Traffic assignment. Modal shares are assigned to the network.

Since there will be only one HSR route determined, this study will use only steps 1 to 3 to evaluate transportation demand on each of feasible routes to produce one optimal route in the end.

TDA can be used either in classical way, or in an integrated way, when steps 1 and 2 are integrated into single demand model. Choice of the approach depends on the data available and the case analyzed. For the case of this study the integrated approach similar to the one, described in Ehrenbreger *et al*. (2010), is used.

The gravity model for TDA will use such data

- left side of the equation:
  - P_i, P_j – cities' population;
  - G_i, G_j – GRP of cities i, j;
  - W_i, W_j – average wages;
  - d_{ij} – distance between the cities i,j.
- right side of the equation:
  - T_{ij} – travel amount on the link between cities i, j.

So the initial equation is:

$$\alpha_0 (P_i P_j)^{\alpha_1} (G_i G_j)^{\alpha_2} (W_i W_j)^{\alpha_3} \left(\frac{1}{d_{ij}}\right)^{\alpha_4} = T_{ij} \qquad (15)$$

Of course, while running the regression, the model can alter, because some hidden correlations between the independent variables can be revealed, or some variables will be statistically not significant.

3.5 Construction complexity assessment

Evaluation of HSR routes depends strongly on the socio-economic factors, but engineering issues should also be taken into consideration. The construction complexity directly influences the final cost of the project, so it should be one of the criteria for final route evaluation, along with travel demand. It has to be measured in the formal way to be included into TOPSIS.

The construction complexity components can be measured either using GIS or by analyzing large-scale topographic maps that provide information about relief (using contour lines), water bodies, cities and communities, roads etc. The approach is more general than that is applied by Ehrenbreger *et al*. (2010), because there is no data to make assumptions about basic construction cost and coefficients to multiply.

The method used can be described as follows:

1. Each the route is traced upon the topographical map.
2. It is approximately adjusted to the real terrain.
3. The terrain conditions such as bad slopes, built-up areas crossings and water body crossings are studied along the route.

Finally, each route is characterized by penalty score, so that the smallest penalty is, the less construction complexity is.

3.6 External effects evaluation

External effects, caused by HSR construction can be evaluated by the same approach as for construction complexity. For this purpose data about living area and protected natural zone crossing is measured using topographic map in a GIS.

IV. Case Study

4.1 Alternative generation

4.1.1 Route generation

The case study of HSR line in Ukraine is conducted here using the planning method proposed in the previous chapter.

A specific application (Route Planner) was designed to implement the algorithm hereinabove described. Usually, different kinds of C programming language are used to create PC software (such as C, C++, C#), but in this study Visual Pascal will be used because of author's better knowledge of it. The development software used is Borland Delphi 7 Lite (2002).

While designing an application, two alternative approaches where initially considered:

- application can be made in an easy style with a purpose to handle just the case of this study;
- application could be made flexible, to be used with different assumptions and for different cases, but it requires more design efforts.

Finally, the decision was to use a flexible approach: user of the application is allowed to load different input files and set up different assumptions for each case.

The input file for Route Planner contains this data:

- Graphical map of the country
- Data about cities locations to be placed on the graphical map
- Data about distances among the cities.

As generation of this file requires large amount of calculations (calculate distances, transform city locations), a simple auxiliary application (iMap Creator) was designed. It takes graphical map and data about city coordinates and transforms them into the input file for Route Planner.

At the screenshots (Figure 10 and Figure 11), the screen of an iMap Creator is illustrated. After setting up a correspondence between two cities' geographical and display coordinates, all necessary calculations are performed and the obtained data is saved into the file, required by Route Planner.

Figure 10 Cities, located at the map in the iMap Creator

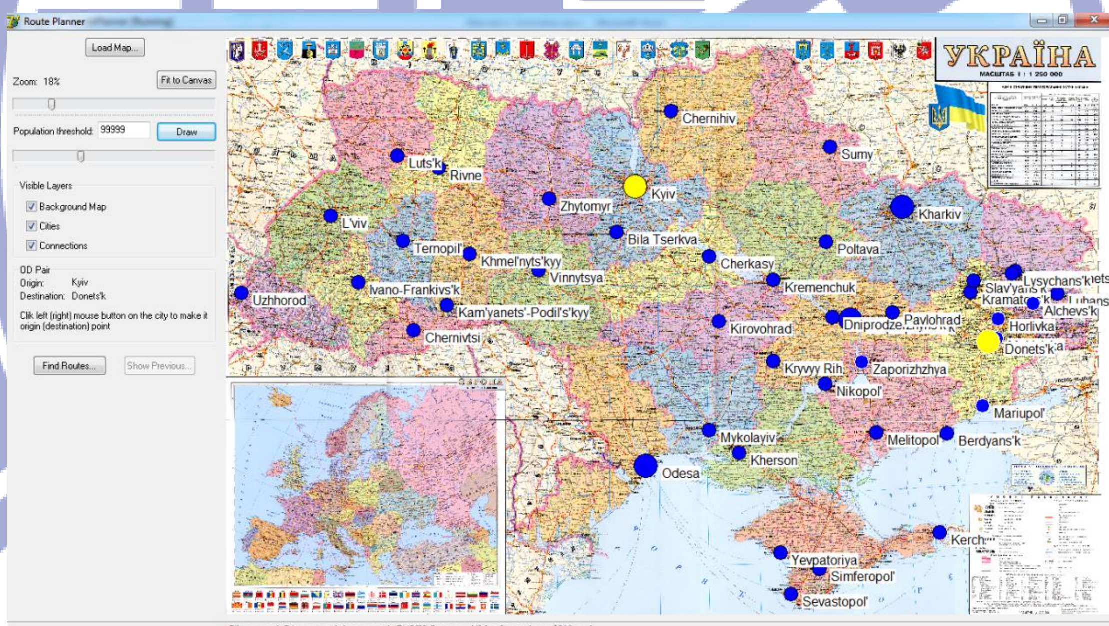
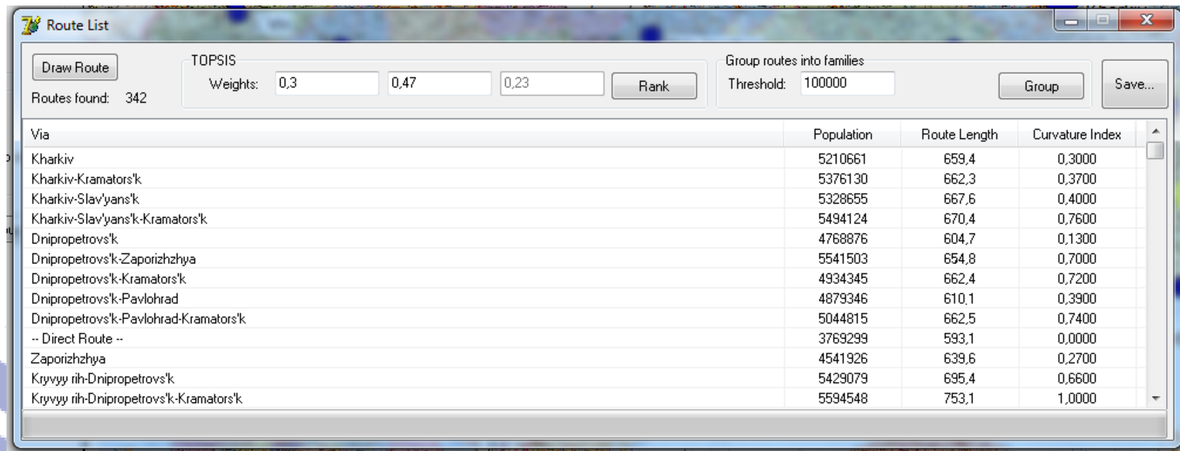


Figure 11 Map, opened in the Route Planner

After the input file is opened in the Route Planner, a population threshold is set up: the algorithm can consider all cities, available for analysis (almost 300), but most of them are small and irrelevant for HSR planning, while their analysis will dramatically increase computing complexity and time sent for computations.

The decision about threshold value is made with respect to the population of the smallest regional center of Ukraine (Uzhhorod, 116 556), so the value of 100 000 as threshold is considered to be reasonable. As this study is devoted to the route between Kyiv and Donetsk, these cities are set up to be origin and destination respectively. Finally, the computation is launched.



Via	Population	Route Length	Curvature Index
Kharkiv	5210661	659,4	0,3000
Kharkiv-Kramators'k	5376130	662,3	0,3700
Kharkiv-Slav'yans'k	5328655	667,6	0,4000
Kharkiv-Slav'yans'k-Kramators'k	5494124	670,4	0,7600
Dnipropetrovs'k	4768876	604,7	0,1300
Dnipropetrovs'k-Zaporizhzhya	5541503	654,8	0,7000
Dnipropetrovs'k-Kramators'k	4934345	662,4	0,7200
Dnipropetrovs'k-Pavlohrad	4879346	610,1	0,3900
Dnipropetrovs'k-Pavlohrad-Kramators'k	5044815	662,5	0,7400
-- Direct Route --	3769299	593,1	0,0000
Zaporizhzhya	4541926	639,6	0,2700
Kryvyy rih-Dnipropetrovs'k	5429079	695,4	0,6600
Kryvyy rih-Dnipropetrovs'k-Kramators'k	5594548	753,1	1,0000

Figure 12 Route list window

The result of algorithm computations – window with full list of all possible 342 routes is provided on the screenshot (Figure 12). This window also provides route ranking that is described later. Visual representation of the list (Figure 13) demonstrates visual reasonability of the algorithm, because the majority of the routes lie in the feasible area.

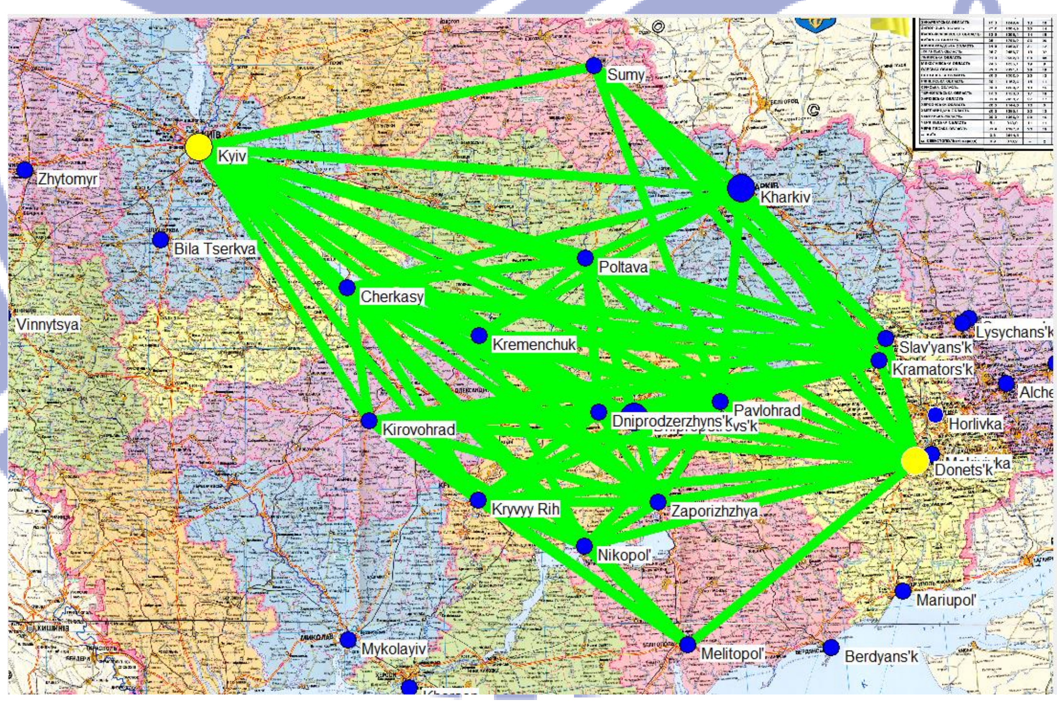


Figure 13 Visual representation of all possible routes

4.1.2 Route ranking

After full list of available routes was generated, a shortlist of candidate routes is created. This process involves following steps:

1. Gathering information about weights of criteria
2. Ranking the routes according to their scores.

Each route is characterized by 3 criteria: route population, route length and route curvature (RP, RL and RC respectively). To obtain the weights of criteria, AHP survey was done, the sample of AHP questionnaire can be found in the Attachment 1.

Four experts were asked to fill in the AHP questionnaires:

- C
  - Professor, Institute of Traffic and Transportation, National Chiao-Tung University, Taiwan
  - Ph.D., major in transportation
- F
  - Professor, Institute of Traffic and Transportation, National Chiao-Tung University, Taiwan
  - PhD, major in transportation
- H
  - Professor, Institute of Traffic and Transportation, National Chiao-Tung University, Taiwan
  - Ph.D., major in management science
- H1
  - Professor, Institute of Traffic and Transportation, National Chiao-Tung University, Taiwan
  - Ph.D., major in transportation

The survey revealed several shortcomings of the criteria, selected for candidate route selection. Mostly, criticism of experts included two elements of the approach:

1) **The criterion of route curvature is too much dependent on the criterion of route length.** Indeed, the more times the route changes its direction, the longer it becomes. Combining this with a fact that optimization direction for both criteria is the same, it is possible to state that route curvature is really redundant.

2) **This approach is too simple and does not include another very significant socio-economic and technical data about the route.** This refers to the initial tradeoff of two-step approach, developed in this study: traditional transportation planning approach can utilize entire amount of data, that influences transportation demand, but its implementation is huge and requires strong mathematical and statistical efforts to provide results. This limits application of

the classical approach to the case, when there is strictly limited amount of alternatives to be evaluated. On the other hand, there is no alternatives initially provided for this study and usage of traditional approach will be extremely complicated for thousands of possible routes.

Technically it is possible to add more raw data to the alternative generation algorithm, such as terrain information or socio-economic information, but this can strongly increase algorithm complexity. In addition to this, the influence of socio-economic characteristics on travel demand is not straightforward enough, to use them simply as criteria for ranking the list of routes. As this data is very important for decision-making, it will be used on the second stage of the analysis: when candidate routes will be shortlisted.

So, taking into account the opinion of experts, only information about weights of RP and RL criteria from AHP survey data will be considered in the further analysis, while RC data will be rejected as redundant.

According to AHP surveys, 3 opinions (one of the experts did not fill in the questionnaire) about weights are available as in Table 7.

Table 7 Weights of criteria

#	RP	RL
1	0.86	0.14
2	0.5	0.5
3	0.6	0.4
Average	**0.65**	**0.35**

Thus, routes in the full list, obtained via Alternative generation algorithm are ranked using TOPSIS according to the criteria of route population and route Length, with average wages of 0.65 and 0.35 respectively.

As Route Planner was initially designed to handle 3 criteria and the criterion of route curvature was rejected, the field for RC is filled with zero value as shown on Figure 14.

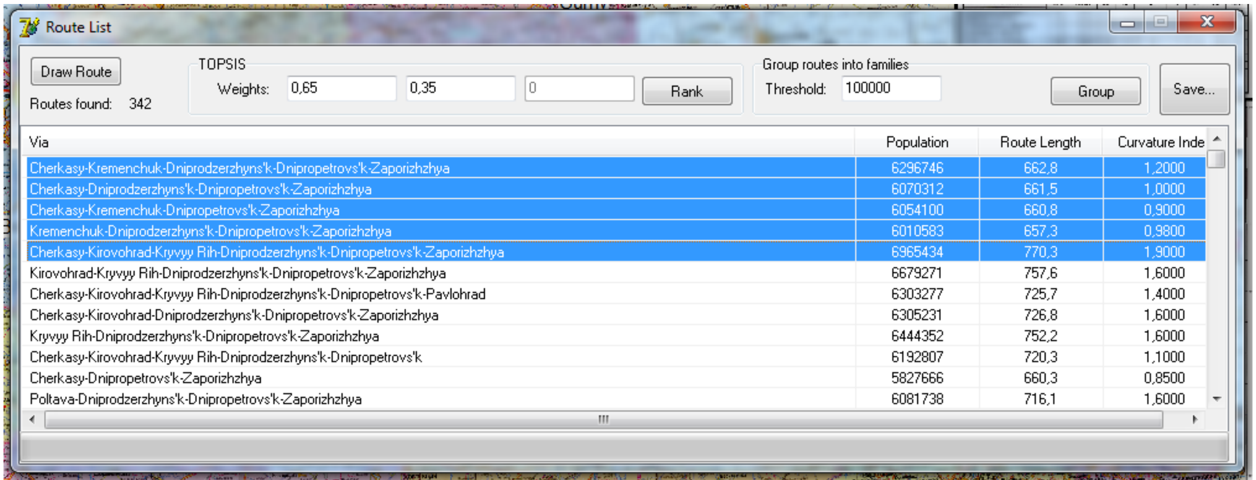


Figure 14 Ranked list of routes, top 5 selected

On this stage an issue that has not been predicted was found. It appeared, that 4 of the top-5 routes, that are intended to be used in future analysis are almost one the same route with minor differences (Table 8 and Figure 15, please note, that lines represent only the connections and do not take into account real terrain). So, the routes have to be grouped into families of almost-the-same routes, where the best route of the family represents it in the evaluation.

Table 8 Route alternatives

	Route Via	Population	Length, km
	Cherkasy-Kremenchuk-Dniprodzerzhyns'k-Dnipropetrovs'k-Zaporizhzhya (**solid black**)	6296746	662.8
	Cherkasy-Dniprodzerzhyns'k-Dnipropetrovs'k-Zaporizhzhya (**dashed black**)	6070312	661.5
	Cherkasy-Kremenchuk-Dnipropetrovs'k-Zaporizhzhya (**grey**)	6054100	660.8
	Kremenchuk-Dniprodzerzhyns'k-Dnipropetrovs'k-Zaporizhzhya (**solid white**)	6010583	657.3
	Cherkasy-Kirovohrad-Kryvyy Rih-Dniprodzerzhyns'k-Dnipropetrovs'k-Zaporizhzhya (**dashed white**)	6965434	770.3

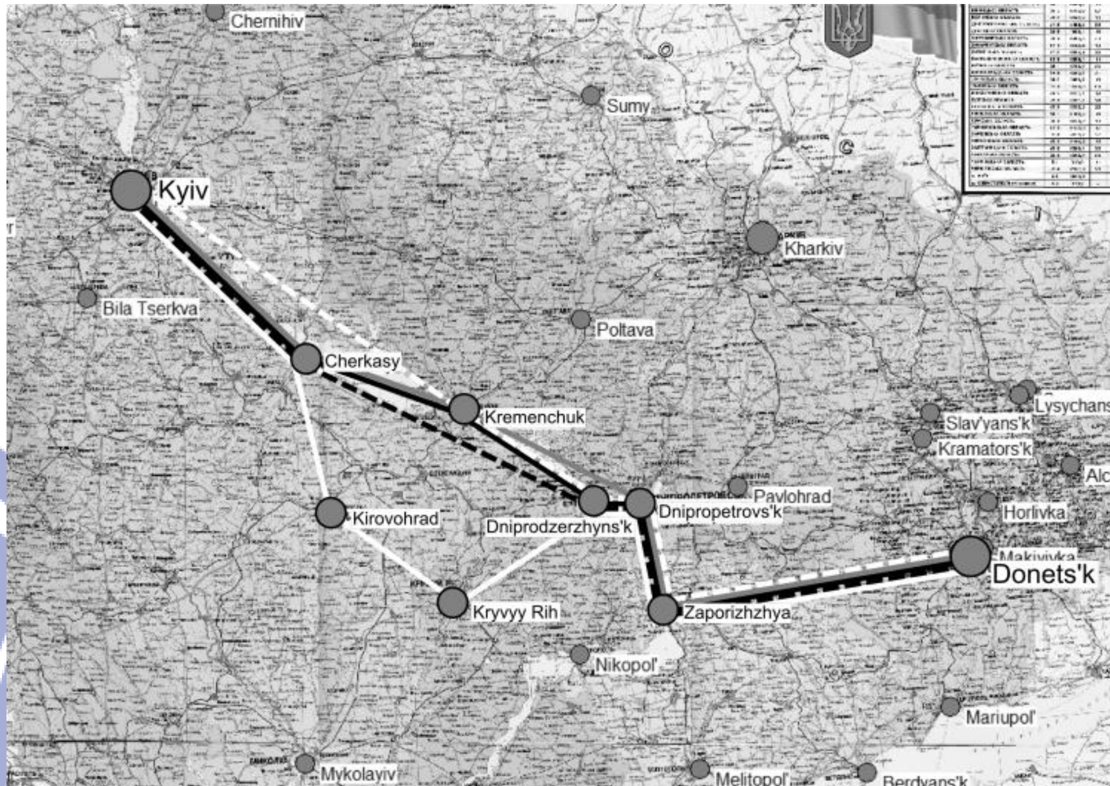


Figure 15 Visual representation of top-5 routes, without grouping

4.1.3 Route grouping

The aim of this part is to define, how to distinguish route families. The key idea of the algorithm, proposed to do so, is to measure the scope of likeliness between two routes by measuring square of the polygon, created by two routes that are being compared. Formally, algorithm goes like this:

1. The first route in the list is taken as a family leader.
2. Each next route is compared to the family leader, if the polygon square is less than that threshold, the route is considered to be a member of current family and excluded from the route list.
3. Return to 1 until there are routes in the list that do not have a family assigned.

There is no way to get reasonable threshold value, but to study its influence on the result empirically. Finally, the value t = 100000 was chosen for grouping.

The result after route grouping is a table, containing 18 candidate routes. It can be noticed, that the fact that RP > RL strongly influenced the resulting list: the top-5 (see Table 9) is mostly oriented on population coverage, while the length of the route is mostly larger than the average of the full list of 342 routes (average is 702.7 km).

Table 9 List of the alternatives, after grouping

#	Route Via	Population	Length, km
1	Cherkasy-Kremenchuk-Dniprodzerzhyns'k-Dnipropetrovs'k-Zaporizhzhya	6296746	662.8
2	Cherkasy-Kirovohrad-Kryvyy Rih-Dniprodzerzhyns'k-Dnipropetrovs'k-Zaporizhzhya	6965434	770.3
3	Cherkasy-Kirovohrad-Kryvyy Rih-Dniprodzerzhyns'k-Dnipropetrovs'k-Pavlohrad	6303277	725.7
4	Poltava-Dniprodzerzhyns'k-Dnipropetrovs'k-Zaporizhzhya	6081738	716.1
5	Cherkasy-Kremenchuk-Poltava-Kharkiv-Slav'yans'k-Kramators'k	6304310	753

Finally, the routes are checked for consistency: some of the factors are not included into algorithm, so it can produce results, inconsistent with reality. Routes 1, 2, 3 and 5 are consistent, but route 4 is not: it lies on the left side of the Dnipro river and connects the cities that lie either on the left side of the river, or on both of them, except one city - Dniprodzerzhyns'k. This city fully lies on the right bank of the Dnipro, and it is very expensive to cross the river two more times. So the route cannot be included to the final top-5, although it formally belongs to them. The final list of routes is in the Table 10.

Table 10 Final list of the routes

#	Route Via	Population	Length, km
1	Cherkasy-Kremenchuk-Dniprodzerzhyns'k-Dnipropetrovs'k-Zaporizhzhya (**solid black**)	6296746	662.8
2	Cherkasy-Kirovohrad-Kryvyy Rih-Dniprodzerzhyns'k-Dnipropetrovs'k-Zaporizhzhya (**dashed black)**	6965434	770.3
3	Cherkasy-Kirovohrad-Kryvyy Rih-Dniprodzerzhyns'k-Dnipropetrovs'k-Pavlohrad (**grey**)	6303277	725.7
4	Cherkasy-Kremenchuk-Poltava-Kharkiv-Slav'yans'k-Kramators'k (**solid white**)	6304310	753
5	Cherkasy-Kremenchuk-Dniprodzerzhyns'k-Dnipropetrovs'k-Pavlohrad (**dashed white**)	5634589	618.2
6	Poltava-Dnipropetrovs'k-Zaporizhzhya	5839092	698
7	Poltava-Kharkiv-Slav'yans'k-Kramators'k	5791713	693.7
8	Cherkasy-Kremenchuk-Dniprodzerzhyns'k-Dnipropetrovs'k-Kramators'k	5689588	670.4
9	Cherkasy-Kirovohrad-Kryvyy Rih-Nikopol'-Zaporizhzhya	5843211	714.1
10	Sumy-Kharkiv-Slav'yans'k-Kramators'k	5763787	710.9
11	Cherkasy-Kharkiv-Slav'yans'k-Kramators'k	5780287	724.1
12	Kharkiv-Slav'yans'k-Kramators'k	5494124	670.4
13	Poltava-Dniprodzerzhyns'k-Dnipropetrovs'k-Pavlohrad	5419581	671.4

14	Poltava-Dniprodzerzhyns'k-Dnipropetrovs'k-Kramators'k	5474580	723.7
15	Cherkasy-Kirovohrad-Kryvyy Rih-Nikopol'-Melitopol'	5227469	784.9
16	Kirovohrad-Kryvyy Rih-Nikopol'	4784421	697.7
17	Kirovohrad-Kryvyy Rih-Melitopol'	4821306	769.1
18	Sumy	4038962	695.7

Top 5 routes that will be considered in further analysis are illustrated on the Figure 16.

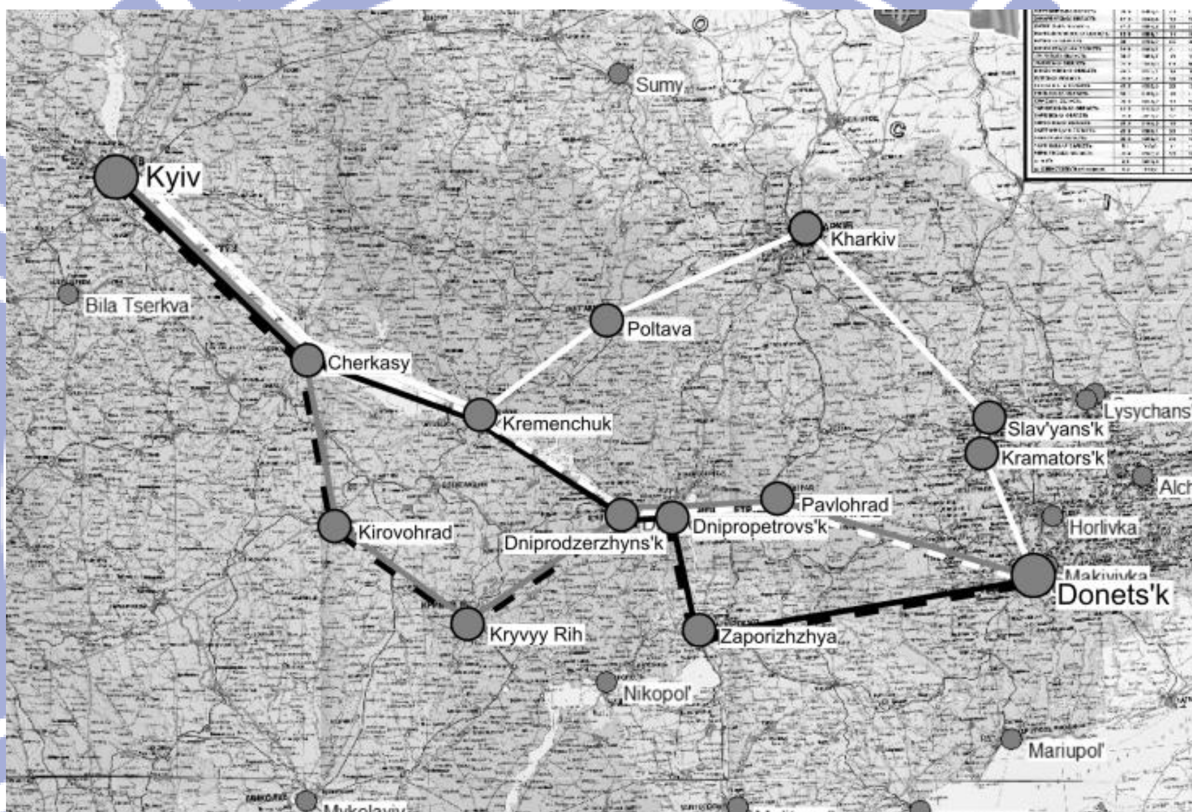


Figure 16 Final top-5 routes for analysis

4.2 Alternative evaluation

4.2.1 Transportation demand analysis

The model that was originally offered for travel demand estimation requires large amount of empirical data. It is usually obtained from statistics, but this study faced several problems concerning the empirical data:

- Socio-economic data published by State Statistics Committee of Ukraine (hereinafter SSCU) is usually categorized by regions, so it is not easy to obtain information about Gross Regional Product and average wages in each city.
- Statistics of transportation in Ukraine is also categorized by regions; moreover it does not provide data about travel amount among the regions: the only published data is the total amount of departures from each region.

The solution to the first issue was found in the annual statistics published by the Association of Ukrainian Cities (hereinafter AUC). Although this organization uses data by SSCU, they query the exact data they need, so the values for individual cities can be found, though the data is still not exactly the one that is required by the model.

The data provided by AUC includes:

- Budget spending including transfers
- Investments
- Consumption of goods
- Average wages

On the other hand, as it can be recalled from Macroeconomics, GDP of a country is

$$Y = C + I + G + NX \qquad (16)$$

where C: private consumption
I: private investments
G: governmental spending
NX: net export

The same formula is used for gross regional product (GRP), but net export (NX) is substituted by the transfers (T):

$$Y = C + I + G + T \qquad (17)$$

So, the data, obtained from AUC can be used to calculate GRP because that data exactly correspond the components of the formula:

- **G + T** is budget spending including transfers

- **I** is investments
- **C** is consumption of goods

The second issue is the lack of data on trips among cities. Although SSCU was queried, they did not provide this data and an alternative approach was needed.

Since city data was not available, this study calibrated travel demand model by regional data. Such alternative could result in underestimations of the demand.

The data, that is obtained from SSCU for calibrating models includes:

- Gross regional product
- Average wages
- Overall departures
- Population

The GRP and average wages data cannot be used "as-is" and they need adjustments: according to the administrative division of Ukraine, two cities (Kyiv and Sevastopol') are excluded from the region they are situated in. This means that statistics of the two regions excludes the data of the corresponding cities. To incorporate them into aggregate data, the conversion was performed as follows:

$$GRP_{KR+} = \frac{P_{KR-}GRP_{KR-} + P_{KC}GRP_{KC}}{P_{KR-} + P_{KC}} \quad (18)$$

where GRP_{KR+}: aggregated GRP of Kyiv Region;
GRP_{KR-}: GRP of Kyiv Region except Kyiv City;
GRP_{KC}: GRP of Kyiv City;
P_{KR-}: population of Kyiv Region except Kyiv City;
P_{KC}: population of Kyiv City.

$$AW_{KR+} = \frac{P_{KR-}AW_{KR-} + P_{KC}AW_{KC}}{P_{KR-} + P_{KC}} \quad (19)$$

where AW_{KR+}: aggregated average wages of Kyiv Region;
AW_{KR-}: average wages of Kyiv Region except Kyiv City;
AW_{KC}: average wages of Kyiv City;
P_{KR-}: population of Kyiv Region except Kyiv City;
P_{KC}: population of Kyiv City.

The same transformation is performed for Sevastopol City and Autonomous Republic of Crimea.

The amount of overall departures is not exactly the data required by the model, because the amount of trips between origin and destination is needed, but Barash et al. (2012) contains a methodology to obtain the required type of data. The key idea is that travels, which depart from each region, are distributed between all the regions proportionally with product of origin and destination population and inverse proportionally to the distance between origin and destination. Mathematically:

$$T_{ij} = k_i \frac{P_i P_j}{d_{ij}} \tag{20}$$

where T_{ij}: trips between origin *i* and destination *j*;
P_i, P_j: population of origin *i* and destination *j* respectively;
d_{ij}: distance between origin *i* and destination *j*;
k_i: coefficient for origin *i*;

Sum of all the trips that begin in the same origin should be equal to the overall departure as follows:

$$D_i = \sum_j T_{ij} = \sum_j k_i \frac{P_i P_j}{d_{ij}} = k_i \sum_j \frac{P_i P_j}{d_{ij}} \tag{21}$$

where D_i: total departures from origin *i*;
P_i, P_j: population of origin *i* and destination *j* respectively;
d_{ij}: distance between origin *i* and destination *j*;
k_i: coefficient for origin *i*;

As origin and destination in this section are regions, the distance d_{ij} is taken as the distance among respective centers of regions. As a consequence, the coefficient can be calculated:

$$k_i = \frac{D_i}{\sum_j \frac{P_i P_j}{d_{ij}}} \tag{22}$$

Finally the matrix of T_{ij} for each OD pair is calculated.

Having T_{ij} calculated and P_i, P_j, G_i, G_j, W_i, W_j, and d_{ij} obtained from statistics, following model can be calibrated in the Excel.

$$\alpha_0 (P_i P_j)^{\alpha_1} (G_i G_j)^{\alpha_2} (W_i W_j)^{\alpha_3} (d_{ij})^{\alpha_4} = T_{ij} \tag{23}$$

where T_{ij}: trips between origin *i* and destination *j*;
P_i, P_j: population of origin *i* and destination *j* respectively;

G_i, G_j: GRP of origin *i* and destination *j* respectively;
W_i, W_j: average wages of origin *i* and destination *j* respectively;
d_{ij}: distance between origin *i* and destination *j*;

As it is a multiplication model, it should be taken into logarithm:

$$\ln(\alpha_0) + \alpha_1 \ln(P_i P_j) + \alpha_2 \ln(G_i G_j) + \alpha_3 \ln(W_i W_j) + \alpha_4 \ln(d_{ij}) = \ln(T_{ij}) \quad (24)$$

The results of the regression analysis are in the Table 11 and Table 12.

Table 11 Regression statistics

Regression Statistics	
Multiple R	0.8785
R Square	0.7717
Adjusted R Square	0.7701
Standard Error	0.5709
Observations	600

Table 12 Coefficiens

	Coefficients	*Standard Error*	*t Stat.*	*P-value*
Intercept	38.5946	3.7634	10.2552	0.0000
ln(GG)	1.5132	0.0619	24.4647	0.0000
ln(WW)	-2.6291	0.3111	-8.4508	0.0000
ln(PP)	-0.8393	0.0874	-9.6023	0.0000
$\ln(d_{ij})$	-0.8312	0.0373	-22.2686	0.0000

As it can be seen, the formal criteria of model applicability show good results, but this model fails the consistency test. In fact, the signs of coefficients conflict with their logic, for example the populations of origin and destination should positively influence the amount of travelling that confronts the nature of the travel demand.

For this reason another model was investigated. After several calibrations, the model was finalized as:

$$\alpha_0 (P_i P_j)^{\alpha_1} (G_j)^{\alpha_2} (W_i)^{\alpha_3} (d_{ij})^{\alpha_4} = T_{ij} \quad (25)$$

$$\ln(\alpha_0) + \alpha_1 \ln(P_i P_j) + \alpha_2 \ln(G_j) + \alpha_3 \ln(W_i) + \alpha_4 \ln(d_{ij}) = \ln(T_{ij}) \quad (26)$$

where T_{ij}: trips between origin *i* and destination *j*;

P_i, P_j: population of origin *i* and destination *j* respectively;
G_j: GRP of destination *j*;
W_i: average wages of origin *i*;
d_{ij}: distance between origin *i* and destination *j*;

The results of the model calibration can be found in Table 13, Table 14, Table and 15.

Table 13 Second regression statistics

Regression Statistics	
Multiple R	0.8656
R Square	0.7493
Adjusted R Square	0.7476
Standard Error	0.5982
Observations	600

Table 14 Second regression statistics (2)

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	4.0000	636.5502	159.1375	444.6408	0.0000
Residual	595.0000	212.9513	0.3579		
Total	599.0000	849.5014			

Table 15 Coefficients of the second regression

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-9.077	1.627	-5.577	0.000
$\ln(W_i)$	0.952	0.327	2.909	0.004
$\ln(G_j)$	0.921	0.034	26.696	0.000
$\ln(P_iP_j)$	0.667	0.097	6.897	0.000
$\ln(d_{ij})$	-0.893	0.039	-22.856	0.000

On the contrary to the original model, this model estimates the coefficients that are consistent with the logic of influence of each independent variable on the dependent one.

According to the formal regression results, final model explains the variation of trip amount between origin and destination at 75% that is a good result for the real statistical data. All the coefficients are statistically significant for both 5% and 1% thresholds. According to the residual analysis, the mean is -2.32792E-14 ≈ 0.

Still, the model is not perfect and this can be seen from the distribution of residuals on the Figure 17.

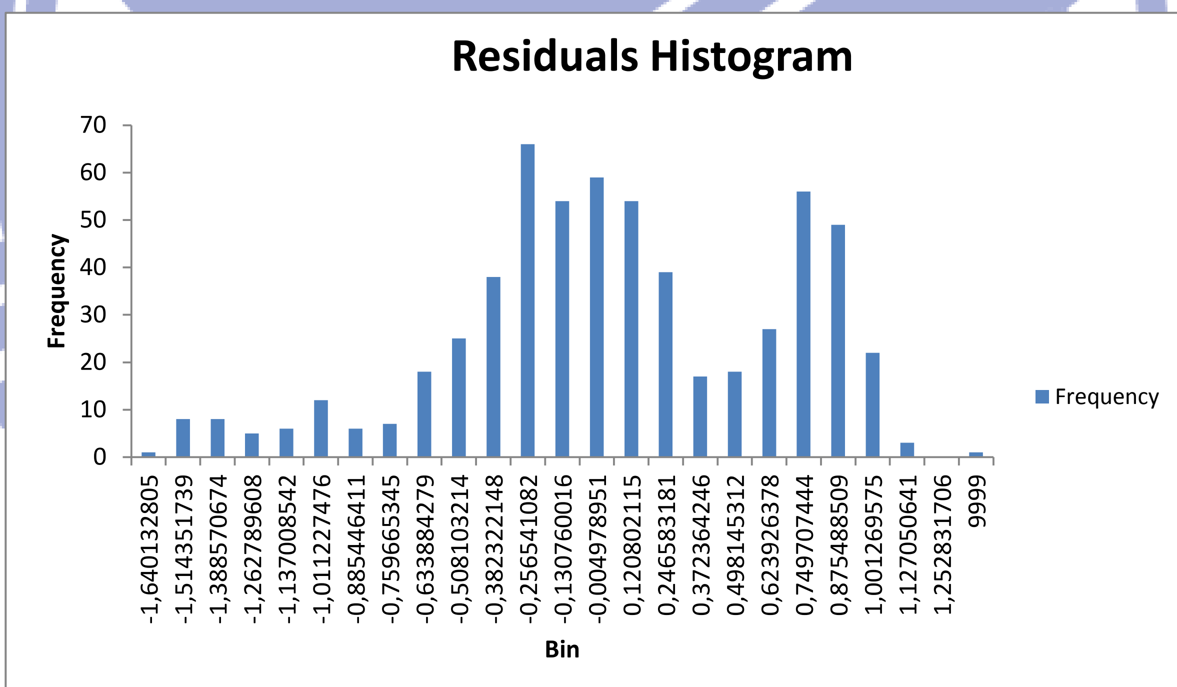


Figure 17 Residuals of the second regression

Formally the normality of residuals' distribution can be tested by χ^2 testing. As Excel data analysis module does not provide any of residual analysis tools, the testing is performed manually, using statistical functions of Excel.

Having all the data classified into bins by the Histogram tool, the following formula is applied:

$$\chi^2 = \sum_i \frac{(E_i - O_i)^2}{E_i} \tag{27}$$

where χ^2: chi-square statistic;
E_i: expected number of samples in the bin *i*;
O_i: observed number of samples in the bin *i*;

Initially, the expected number of samples in each bit is unknown, and it has to be generated according to the formula:

$$E_1 = f(x_1, \mu, \sigma^2) \tag{28}$$

$$E_i = C(f(x_i, \mu, \sigma^2) - f(x_{i-1}, \mu, \sigma^2)), i > 1 \tag{29}$$

where E_i: expected number of samples in the bin *i*;
x_i: upper bound of the bin *i*;
μ: mean of the residuals;
σ^2: variance of the residuals.

The upper bounds of the bins are provided by the Histogram tool, mean and variance on the residuals can be found in the descriptive statistics of residuals (see Table 16).

Table 16 Residuals' statistics

	Residuals
Mean	-2.32792×10^{-14}
Standard Error	0.02434171
Median	-0.046428914
Mode	#N/A
Standard Deviation	0.596247701
Sample Variance	0.355511321
Kurtosis	-0.141485371
Skewness	-0.370276461
Range	3.018745577
Minimum	-1.640132805
Maximum	1.378612772
Sum	-1.39675×10^{-11}
Count	600

Finally, the expected frequencies are calculated, they can be seen in the Table 17.

Table 17 Observed frequencies vs. expected frequencies

Bin	*Obs. Frequency*	*Exp. Frequency*
-1.640132805	1	2
-1.514351739	8	2
-1.388570674	8	3
-1.262789608	5	4
-1.137008542	6	7
-1.011227476	12	10
-0.885446411	6	14
-0.759665345	7	20
-0.633884279	18	26
-0.508103214	25	32
-0.382322148	38	38
-0.256541082	66	44
-0.130760016	54	48
-0.004978951	59	50
0.120802115	54	50
0.246583181	39	48
0.372364246	17	44
0.498145312	18	39
0.623926378	27	32
0.749707444	56	26
0.875488509	49	20
1.001269575	22	15
1.127050641	3	10
1.252831706	0	7
More	1	11

Using this data, chi-square statistics can be calculated and compared to the critical value at α = 0.05:

$$\chi^2 = \sum_i \frac{(E_i - O_i)^2}{E_i} = 201.03 \tag{30}$$

$$\chi^2_{critical}(\alpha; df) = \chi^2_{critical}(0.05;22) = 33.92 \tag{31}$$

where α: confidence level set as 0.05;

df: degrees of freedom, calculated as df = #bins – 1 - #parameters calculated from the sample = 25-1-2 =22.

As $\chi^2 > \chi^2_{critical}$, we can say, that the distribution of residuals is **not normal.**

Although the model was changed in different ways, none of the tested alternatives gave satisfactory results without affecting model consistency.

So, although the residuals distribution is not normal, this model was used to estimate travel demand among route cities. Further research should be conducted in the future to reach further valid and reliable estimations of travel demand.

According to the results of the regression analysis, travel demand for each O-D pair is measured as follows:

$$T_{ij} = e^{-9.077}(P_i P_j)^{0.667}(G_j)^{0.921}(W_i)^{0.952}(d_{ij})^{-0.893} \tag{32}$$

Since originally the model was calibrated for conventional railway demand, T_{ij} is a conventional railway demand too, so modal shift analysis should be performed.

HSR demand is determined according to the following mode shift assumptions:

- If there is no railway link between city i and city j, the full amount of T_{ij} is shifted to HSR
- Otherwise only people that use top-class trains will shift, because they are ready to pay for speed and/or comfort

According to the service schedule of the Ukrainian railroads, the total amount of top-class trains (with average speeds over 60 km/h) is 39.38%.

HSR demand that will be used for route evaluation is calculated according to the following formula

$$HSRD_{ij} = (1 - CR_{ij}(1 - ms))T_{ij} d_{ij} \tag{33}$$

where $HSRD_{ij}$: HSR travel demand on the link between origin *i* and destination *j*;
CR_{ij}: existence of conventional rail between origin *i* and destination *j* (0 or 1);
ms: mode shift level, ms = 39.38%;
CR_{ij}: conventional rail trips on the link between origin *i* and destination *j*;
d_{ij}: distance between origin *i* and destination *j*.

The results of estimated O-D demands ($HSRD_{ij}$) for all routes are listed in the Appendix 3 and the total travel demands for the five routes can be found in the Table 18.

Table 18 Travel demand for each route

Routes	Route 1	Route 2	Route 3	Route 4	Route 5
Estimated demand, pass-km	1 189 890 691	1 584 618 417	1 317 486 515	943 512 618	728 907 272

4.2.2 Construction complexity evaluation

The construction complexity evaluated in this study contains 3 components:

- Water body crossing index
- Built-up area crossing index
- Bad slope index

All the indexes are computed by tracing routes in a GIS using the topographical map and digital elevation data. There are many IT solutions that provide GIS interface. After some trials, ArcGIS was chosen to perform all necessary estimations.

First, water body crossings and built-up area crossings were analyzed using World Topo Map, provided by Esri on the scale of 1:25 000. Although the map is provided in the Mercator projection, it was transformed to the Lambert Azimuthal Equal Area projection to make distance measurement possible in a correct way. Routes are traced on the map according to several rules:

- Route segments among cities should be as straight as possible.
- The Dnipro River cannot be crossed more than once.
- HSR line inside a city that has a HSR station can use the same station existing railway does, where possible.

Figure 18 illustrates Route 5 traced over a topographic map as an example.

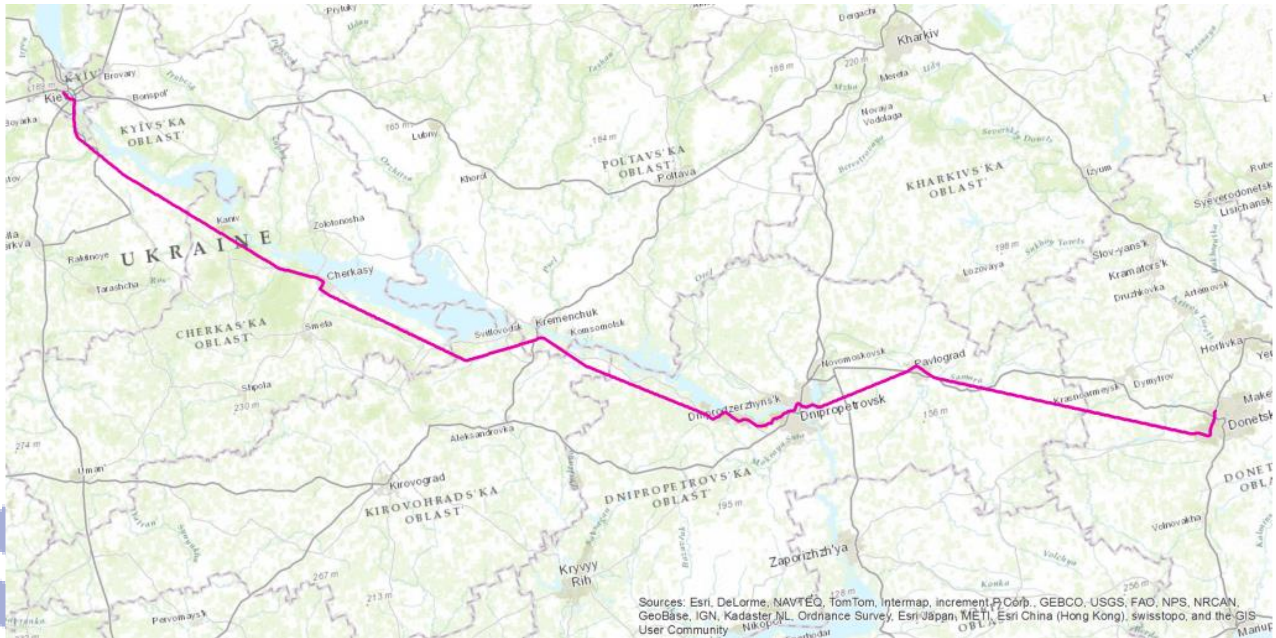


Figure 18 Route 5 traced over the topographic map

As candidate routes have some similar parts, all of them were split into unique sections in order to avoid multiple analyses of the same sections.

The five candidate routes are as follows:

1. Kyiv - Cherkasy - Kremenchuk - Dniprodzerzhyns'k - Dnipropetrovs'k – Zaporizhzhya - Donets’k.
2. Kyiv - Cherkasy - Kirovohrad - Kryvyy Rih - Dniprodzerzhyns'k - Dnipropetrovs'k - Zaporizhzhya - Donets’k.
3. Kyiv - Cherkasy - Kirovohrad - Kryvyy Rih - Dniprodzerzhyns'k - Dnipropetrovs'k - Pavlohrad - Donets’k.
4. Kyiv - Cherkasy - Kremenchuk - Poltava - Kharkiv - Slav'yans'k - Kramators'k - Donets’k.
5. Kyiv - Cherkasy - Kremenchuk - Dniprodzerzhyns'k - Dnipropetrovs'k - Pavlohrad - Donets’k

For each section length of water crossings and living area crossings is computed, the results are represented in Table 19.

Table 19 Lengths of sensitive area crossings at each section

Route section	Built-up area crossing, m			Water body crossing, m
	Rural	Urban	Industrial	
Kyiv - Cherkasy	24661	1624	1560	746
Cherkasy - Kremenchuk	15037	1443	0	2326
Kremenchuk - Dniprodzerzhyns'k	15398	8742	0	180
Dniprodzerzhyns'k - Dnipropetrovs'k	0	1341	0	0
Dnipropetrovs'k - Zaporizhzhya	5771	5615	4130	4214
Zaporizhzhya - Donets'k	7915	0	0	620
Cherkasy - Kirovohrad	11605	14538	0	772
Kirovohrad - Kryvyy Rih	1906	2973	0	1959
Kryvyy Rih - Dniprodzerzhyns'k	13047	5256	0	240
Dnipropetrovs'k - Pavlohrad	4015	4876	0	3020
Pavlohrad - Donets'k	32539	2374	0	1320
Kremenchuk - Poltava	12186	0	0	739
Poltava - Kharkiv	14247	14725	0	1141
Kharkiv - Kramators'k	9628	21147	0	701
Kramators'k - Donets'k	2038	6047	0	80

By combining data for the sections, overall information about each route is obtained as listed in Table 20.

Table 20 Lengths of sensitive area crossings at each route

Route	Built-up area crossing, m			Water crossing, m
	Rural	Urban	Industrial	
Route 1	68782	18765	5690	8086
Route 2	64905	31347	5690	8551
Route 3	87773	32982	1560	8057
Route 4	77797	44986	1560	5733
Route 5	91650	20400	1560	7592

Finally, values of rural, urban and industrial zone crossing are combined into one under an assumption that land acquiring in the latter two is twice more complicated, so that values are multiplied by two. Then the indexes are calculated with respect to the route length as follows:

$$I_{BA} = \frac{L_{RA} + 2L_{UA} + 2L_{IA}}{L} \qquad (34)$$

where I_{BA} is built-up area crossing index;

L_{RA}, L_{UA}, L_{IA} are total lengths of rural, urban and industrial area crossings respectively;

L is a route length.

$$I_{WB} = \frac{L_{WB}}{L} \qquad (35)$$

where I_{WB} is water body crossing index;

L_{WB} is total length of water body crossings;

L is a route length.

The results are shown in the Table 21.

Table 21 Built-up area crossing and water body crossing indexes

Route	Build-up area crossing index	Water body crossing index
Route 1	0.163	0.011
Route 2	0.169	0.010
Route 3	0.206	0.011
Route 4	0.216	0.007
Route 5	0.205	0.012

On the next step, slope along a route is analyzed using DEM-datafile. DEM is a digital elevation model format that stores elevation data. In this study GTOPO30 dataset by Earth Resources Observation and Science (EROS) Center is used. Elevation data combined with semi-transparent topographic map is shown on Figure 19.

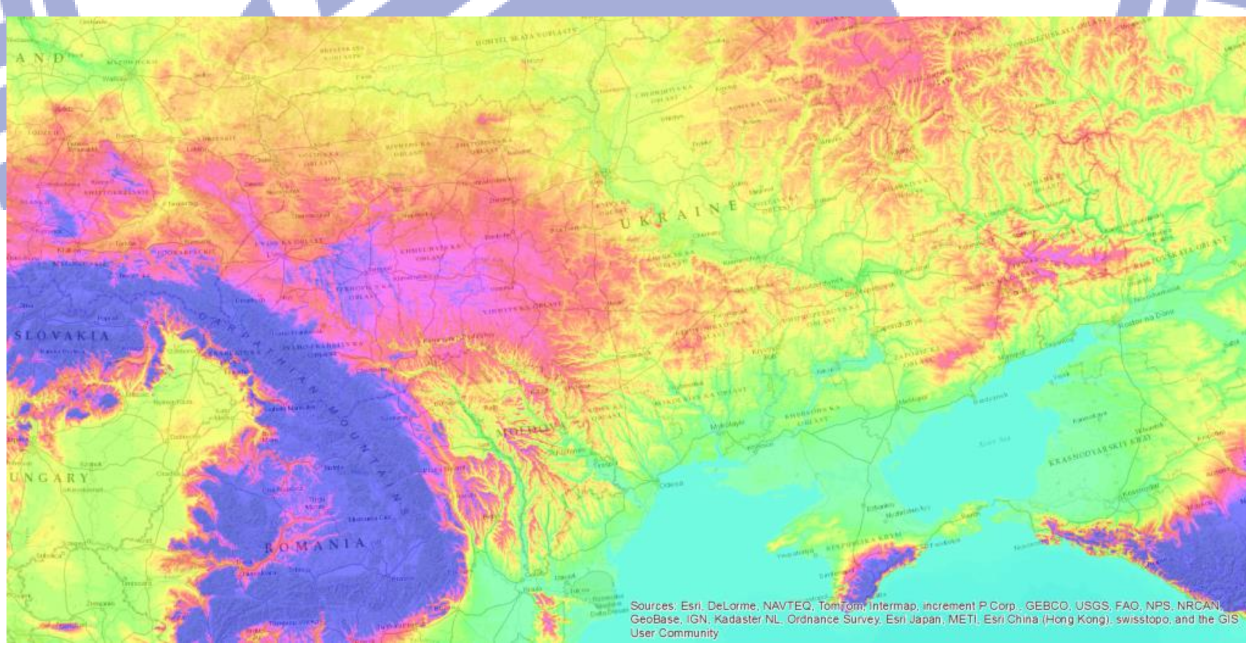


Figure 19 Loaded DEM data in ArcGIS

Initially the routes traced on the basemap do not have z-feature with information about the elevation. To obtain it, each route line is interpolated with the surface DEM-file by Arc GIS (ArcToolbox-> 3D Analyst Tools-> Functional Surface-> Interpolate Shape).

After this action, elevation profile along each route vertex is obtained. Consequently, the slope for each vertex *i* can be calculated as follows:

$$s_i = \frac{|z_i - z_{i-1}|}{x_i - x_{i-1}} \qquad (36)$$

where: s_i – slope of vertex *i*;
z_i – elevation of vertex *i*;
z_i – distance of vertex *i* from the initial one.

Having the list of slope values, the number of bad slopes is counted. Usually critical slope for HSR line is assumed to be 3.5% - 4% (UIC, 2010). In this study the slope is considered to be bad if it exceed 3.5%.

$$I_{BS} = \frac{N(s_i > 0.035)}{L} \qquad (37)$$

where I_{BS}: bad slope index;
s_i: slope value of the route section *i* that lies between vertexes *i-1* and *i* along the route;
L: route length;
N (s_i > 0.035): number of route sections with slopes over the critical value.

The results of the bad slope calculations are in Table 22.

Table 22 Bad slope index

Route	Bad slope number	Bad slope index
Route 1	6	0.0083
Route 2	2	0.0024
Route 3	3	0.0039
Route 4	6	0.0076
Route 5	7	0.0106

So, the components of the construction complexity have the values as shown in Table 23.

Table 23 Construction complexity summary

Route	Built-up area crossing index	Water body crossing index	Bad slope index
Route 1:	0.163	0.011	0.0083

Route 2:	0.169	0.010	0.0024
Route 3:	0.206	0.011	0.0039
Route 4:	0.216	0.007	0.0076
Route 5:	0.205	0.012	0.0106

4.2.3 External effects evaluation

The external effects analyzed in this study contain two perspectives: influence on the living areas and protected natural zones. They are evaluated in a way, similar to the construction complexity.

The first external effect – influence on living areas – partially utilizes the results of analysis that was performed in the previous stage. While built-up areas crossing index was calculated by 3 components: rural, urban and industrial area crossing distances, living area crossing index requires only the first two of them.

The second external effect that concerns influence on the protected areas was evaluated in a less precise way, because World Topo Map, used as a background in ArcGIS, do not provide such kind of information. First the paper topographic maps were studied for every region and then each protected area was found in the GIS, by which the distance of crossing was measured. The calculations are done as follows:

$$I_{LA} = \frac{L_{RA} + 2L_{UA}}{L} \qquad (38)$$

where I_{LA} is living area crossing index;
L_{RA}, L_{UA} are total lengths of rural and urban area crossings respectively;
L is a route length.

$$I_{PNZ} = \frac{L_{PNZ}}{L} \qquad (39)$$

where I_{PNZ} is protected natural zone crossing index;
L_{PNZ} is total length of protected natural zone crossings;
L is a route length.

The result of these estimations is provided in Table 24.

Table 24 External effects summary

Route	Living area crossing index	Protected natural zone crossing index
Route 1:	0.147	0.000
Route 2:	0.155	0.000
Route 3:	0.202	0.000
Route 4:	0.212	0.017
Route 5:	0.201	0.000

It is important to mention the similarity between **built-up area crossing index** and **living area crossing index:** the only difference between them is the component of length of industrial area crossing. This can lead to the correlation of these values, so that it will cause the violation of one of the AHP assumptions: the criteria should be independent.

4.2.4 Decision about the optimal route

This section presents the evaluation results of route planning based on the criteria. The values of criteria for all alternative routes are listed in Table 25.

Table 25 Criteria values for the alternatives

Criteria	Route 1	Route 2	Route 3	Route 4	Route 5
Estimated demand, pass-km	1 189 890 691	1 584 618 417	1 317 486 515	943 512 618	728 907 272
Construction Complexity					
Built-up areas index	0.163	0.169	0.206	0.216	0.205
Water body crossing index	0.011	0.01	0.011	0.007	0.012
Bad slope index	0.0083	0.0024	0.0039	0.0076	0.0106
External effects					
Living area crossing index	0.147	0.155	0.202	0.212	0.201
Protected natural zone crossing index	0	0	0	0.017	0

The decision is based on multiple criteria that refer to different objectives with different optimization direction. They are:

- **Maximize** estimated travel demand (ETD) along the route, criterion:
    - estimated travel demand (TD)
- **Minimize** construction complexity (CC), criteria:
    - built-up area crossing index (BAI)
    - water body crossing index (WBI)
    - bad slope index (BSI)
- **Minimize** external effects (EE), criteria:
    - living area crossing index (LAI)
    - protected natural zone crossing index (PNZI)

TOPSIS was used to rank the alternative routes. This method requires weights of the criteria, so AHP survey was performed first; the questionnaire is included in the Appendix **. In this case three groups of pair wise comparisons were required: for construction complexity, external effects and objectives. Construction complexity and objectives have more than two items, so the weights have to be calculated using AHP method.

Four experts where asked to fill in the survey:

- C
  - Professor, Institute of Traffic and Transportation, National Chiao-Tung University, Taiwan
  - Ph.D., major in transportation
- F
  - Professor, Institute of Traffic and Transportation, National Chiao-Tung University, Taiwan
  - PhD, major in transportation
- L
  - Professor, National Taiwan University, Department of Geography, Taiwan
  - Ph.D., major in transportation
- W
  - Professor, Institute of Traffic and Transportation, National Chiao-Tung University, Taiwan
  - Ph.D., major in transportation

The weights, counted from the pairwise comparisons and the consistency rates of the AHP calculations are provided in the Table 26. All the weights are consistent, so they can be used for the decision-making.

Table 26 Weights of criteria

Criteria	C	F	L	W
Estimated travel demand, pass-km	0.59	0.63	0.63	0.44
Construction Complexity	0.09	0.26	0.11	0.11
Built-up areas index	0.57	0.6	0.26	0.65
Water body crossing index	0.29	0.2	0.11	0.23
Bad slope index	0.14	0.2	0.63	0.12
Consistency ratio	***0***	***0***	***0.033***	***0.003***
External effects	0.32	0.11	0.26	0.44
Living area crossing index	0.5	0.75	0.75	0.5
Protected natural zone crossing index	0.5	0.25	0.25	0.5
Consistency ratio	***0.008***	***0.033***	***0.033***	***0***

Integrated weights are required for TOPSIS, so each criterion weight was multiplied to the weight of corresponding objective (if the objective has more than one criterion). Finally, the weights given by different experts were averaged as it can be seen in Table 27.

Table 27 Integrated weights

Criteria	C	F	L	W	AVG
Estimated demand, pass-km	0.587	0.633	0.633	0.444	0.575
Built-up areas index	0.051	0.156	0.028	0.072	0.077
Water body crossing index	0.026	0.052	0.011	0.026	0.029
Bad slope index	0.013	0.052	0.067	0.014	0.036
Living area crossing index	0.162	0.08	0.195	0.222	0.165
Protected natural zone crossing index	0.162	0.027	0.065	0.222	0.119

Using these average values, alternatives were ranked using TOPSIS, see Table 28.

Table 28 Route rankings

Routes	Rank
Route 2	0.956
Route 3	0.637
Route 1	0.586
Route 4	0.234
Route 5	0.171

So, the Route 2 (see Figure 20) is the best, according to the chosen criteria and their weights.

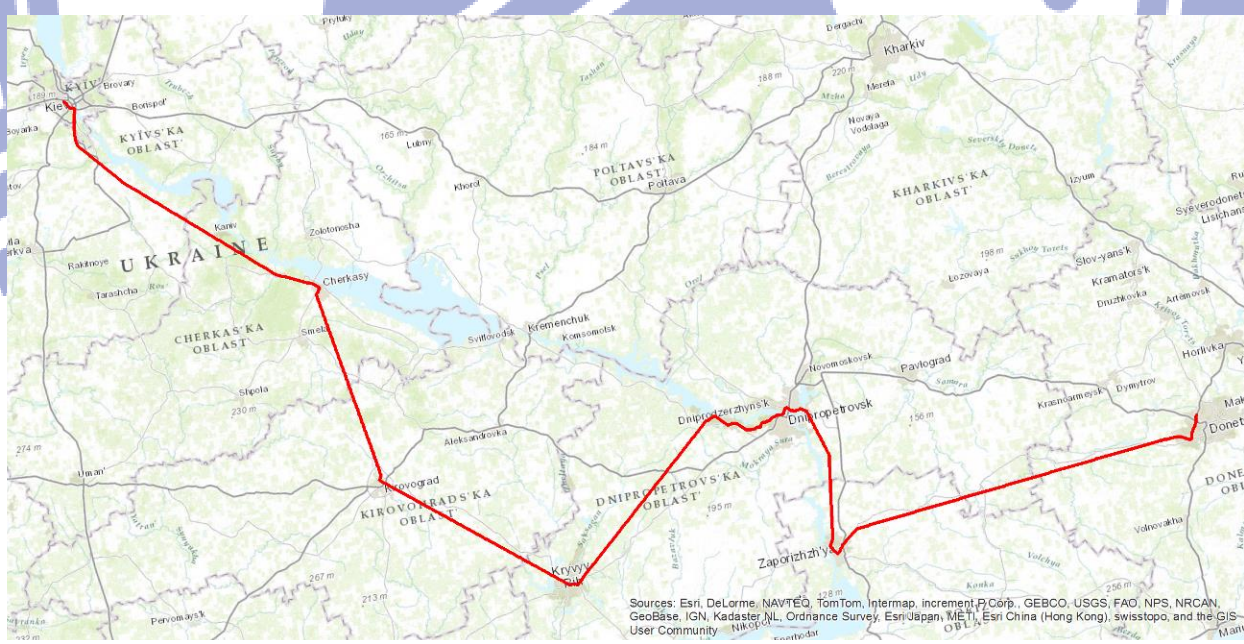


Figure 20 The optimal route

V. Conclusions and Recommendations

This study is a first attempt to plan a HSR route for Ukraine. Since there is no previous academic study or governmental document that can provide information about route alternatives, this study began with the work of alternative generation.

High speed rail studies usually use ready alternatives, so there was no information found about generation of alternative routes. The previous studies usually regard either ready alternatives, or given corridors. For this reason this study introduces a formal algorithm that generates all possible alternative routes using given origin, destination and data about other cities of the country. Instead of spatial limitations (e. g. a predefined corridor), this algorithm uses a set of formal rules that work as constraints. They are applied at the search stage, when the algorithm searches for the cities that can be included into the route from given origin to destination. These rules help to cut off all the routes that cannot be feasible under any conditions.

Although algorithm currently operates only with population and location of cities, it can be improved by adding additional data, such as elevation, or constraints that make links between some cities impossible. This can be useful, for example, when there is a pair of cities on the route lie on the different sides of the wide bay that cannot be crossed. On the other hand, the algorithm should be improved carefully, because it uses recursive search component that can dramatically increase computation time with the increase of amount of the input data and/or constraints. This algorithm is implemented as a software application that was exclusively designed for this study in Borland Delphi 7 Lite.

The routes generated by the application are ranked using TOPSIS with regard to the route length and population coverage. The latter one is assumed to be a proxy to the transportation demand, generated along the line. Using these two criteria, the best five alternatives are chosen from the full list. Because of the weights obtained from the experts, the top-5 routes are mostly concerned about the overall population of the cities, connected by the HSR line.

Final decision about the most optimal route is made with respect to the estimated travel demand, construction complexity (built-up area crossing, water body crossing and bad slopes) and external effects (living area crossing and protected natural zone crossing). The optimal route determined in this study is the route "Kyiv - Cherkasy - Kirovohrad - Kryvyy Rih - Dniprodzerzhyns'k - Dnipropetrovs'k - Zaporizhzhya - Donets'k" with total demand of 1 584 618 417 pass-km.

The best and the second-best rotes are very close to each other, the difference is in one city: instead of Zaporizhzhya the route goes via Pavlohrad. This route is 44.6 km shorter, but Pavlohrad is significantly smaller than Zaporizhzhya (110 470 and 772 627 respectively), so the predicted transportation demand for the second-best route is 1 317 486 515 pass-km, that is less than for the best choice. In the same time it has worse values of computation complexity and external effects indexes.

The scope of this study is limited, because HSR planning process is very complicated, time and recourse consuming. Consequently, this leads to the fact that future research can contribute much to HSR studies in Ukraine. Possible future research directions could be as follows:

- More sophisticated approach to the route planning that includes more objectives that should be satisfied (such as political feasibility, for example).
- As it was already mentioned, this study relies on the AHP criteria that are probably not independent, so it will be useful for the future studies to improve this study by applying criteria that will fully satisfy the assumptions of AHP.
- Route planning with better demand model, based on yearly transportation data between city pairs.
- HSR influence on regional development of Ukraine.
- Project financing study.

Conducting each of these studies requires large amounts of statistical data, issued by the national statistics, so the way statistics is collected and published should be altered. Currently it produces the data that can be used only under certain assumptions that lead to the distortion of results. There are three major improvements that are necessary for the future of the transportation studies in Ukraine:

- Socio-economic data should not only be collected for the regions, but for the most important cities at least.
- Data about transportation amount between the most important cities is also critical. Ukrainian Railroad uses an electronic system, so this data exists, it just has to be processed and published (or provided on demand of the researcher).
- Ukraine needs a national geospatial database that will contain data about elevation, terrain conditions and land use. Currently international data is used, where possible, but it is often issued by either country for itself, or by the large organization for its members. This means that the majority of digital maps for Europe concern only European Union and though omit Ukraine.

So, analysis of possibility of introducing HSR network in Ukraine should be regarded as a long run strategy for both government and academics. This study is just the first step and the amount of future works required to evaluate HSR is huge. One of the most important outcomes of this research is an identification of data shortage, because data that is available at the moment do not allow making a precise and trustworthy route characteristics estimation. That is why, the government of Ukraine should improve the way the national statistics is collected and published to meet the needs of future transportation research.

References

1. Behrens, C. and Pels, E. "Intermodal Competition in the London–Paris Passenger Market: High-Speed Rail and Air Transport", Journal of Urban Economics 17, 278–288, 2012.

2. Barash, Y. S., Charkina, T. Y., Melyantsova, Y. P., Karas', O. O. "Principles for Determining the Efficiency Of Suburban Passenger Trains Operation at a Given Direction", Journal of Dnipropetrovs'k National Railroad Transportation University 41, 234 – 249, 2012.

3. Bierlaire, M. "Mathematical Models for Transportation Demand Analysis", Facultes Universitaires Notre-Dame de la Paix de Namur, Ph.D. Degree in Science, 1995.

4. California High-Speed Rail Authority, Final California High-Speed Train Project Environmental Impact Report/Environmental Impact Statement, 2012.

5. Cascetta, E. and Coppola, P. "High Speed Rail Demand: Empirical and Modeling Evidences from Italy", European Transport Conference 2011, Glasgow, Scotland, UK, 2011.

6. Caulfield, B. and Ghosh, B. "Transport Modelling (Lecture Notes)", 2011 URL: http://www.tcd.ie/civileng/Staff/Brian.Caulfield/T2%20-%20Transport%20Modelling/, download date: May. 15, 2013.

7. CIA World Factbook. URL: https://www.cia.gov/library/publications/the-world-factbook/, download date: Sep. 20, 2012.

8. Dobruszkes, F. "High-Speed Rail and Air Transport Competition in Western Europe: a Supply-Oriented Perspective", Transport Policy 18, 870–879, 2011.

9. Ehrenberger, S., Winter J., Malik. F. "Quantifying the Potentials of a New High Speed Train Using a Gravity Model and GIS", European Transport Conference 2010, Glasgow, Scotland, UK, October 2010.

10. Feng, C. M. "Transportation Project Evaluation", unpublished lecture notes.

11. Hensher, D. "A Practical Approach to Identifying the Market Potential for High Speed Rail: a Case Study in the Sydney-Canberra Corridor", Transportation Research A 31, 431-446, 1997.

12. Ryder, A. "High Speed Rail", Journal of Transport Geography 22 303–305, 2012.

13. UIC. "High Speed Rail: Fast Track to Sustainable Mobility", 2010. URL: http://www.uic.org/IMG/pdf/20101124_uic_brochure_high_speed.pdf

14. Uršej, Š. And Kontić, B. "The Role of Surface Characteristics in Directing Subsurface Spatial Planning Processes: the Case Study of a High-Speed Railway in Slovenia", Tunneling and Underground Space Technology 22, 414–432, 2007.

15. van Wee, B., van den Brink, R. and Nijland, H. "Environmental Impacts of High-Speed Rail Links in Cost–Benefit Analyses: a Case Study of the Dutch Zuider Zee Line", Transportation Research Part D 8, 299–314, 2008.

16. Yao, E. and Morikawa, T. "A Study on Integrated Intercity Travel Demand Model", 10th International Conference on Travel Behaviour Research, Lucerne, 2003.

Appendix 1 AHP questionnaire #1

Dear Sir/Madam,

My name is Anton Hagen and I am doing my master thesis research on the topic “Planning a High Speed Rail Route in an Emerging Country: A Case Study of Ukraine”. In this study I’m going to develop the optimal *high speed rail* (HSR) route between two major cities in Ukraine. Almost all previous HSR studies concern about evaluation of given alternatives, but for this case no alternative is available yet.

This study develops the optimal HSR route via two major stages: alternative generation and alternative evaluation.

At the first stage, all possible routes from origin to destination are identified by the algorithm that uses data of city population and distances among cities, and numerous constraints. Those routes are ranked according to three criteria and the top five routes are considered to be alternatives for further evaluation.

At the second stage more detailed analysis will be provided for each alternative, previously generated. It will include demand modeling, evaluation of engineering complexity and external effects. Finally, the decision about optimal route will be made.

This questionnaire is used at the first stage to choose 5 candidate routes from full route list.

The criteria used for route ranking are:

- **Route population (*RP*)**, which is a sum of populations of cities connected by a route.

Let p_i (i = 1...N) denotes the population of city i along a route, i=1 represents the origin city and i=N means the destination city, then:

$$RP = \sum_{i=1}^{N} p_i \text{ (optimization direction: MAX)}$$

- **Route length (*RL*)**, which is a sum of distances between two sequential cities along a route.

Let d_{ij} (i, j = 1...N) denotes the distance between cities i and j along a route, then

$$RL = \sum_{i=1}^{N-1} d_{i,i+1} \text{ (optimization direction: MIN)}$$

- **Route curvature (*RC*)**, which is a sum of inverse angles between two sequential links.

Let $l_{i,i+1}$ denotes a link connecting two sequential cities i and i+1 and π is a mathematical constant (π = 3.141592…), then

$$\alpha_1 = \pi$$

$$\alpha_i = \angle(l_{i-1,i}, l_{i,i+1}), i = 2...(N-1)$$

$$RC = \sum_{i=1}^{N-1}(\pi - \alpha_i) \text{ (optimization direction: MIN)}$$

The Figure 1 presents a simple example of the above criteria.

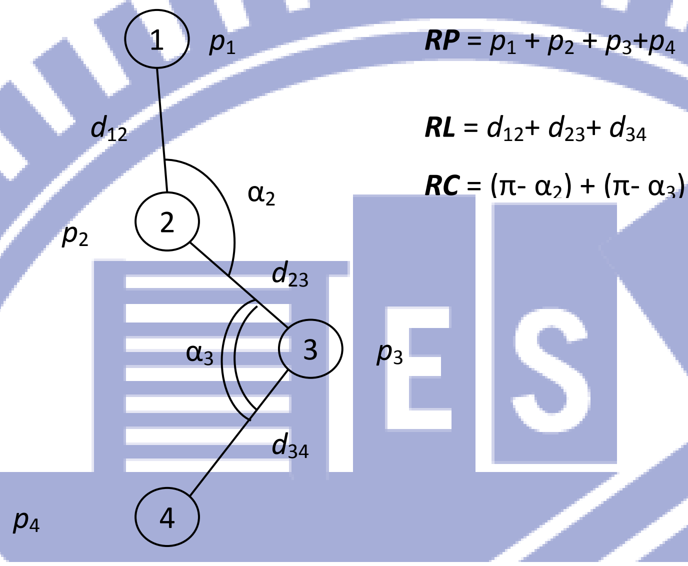


Figure 21. Simple example of ranking criteria

To rank all possible routes according to the criteria of *RP*, *RL* and *RC*, the method of TOPSIS is chosen as the ranking method and it requires weights of the criteria.

Please help me with my thesis and answer the following questions.

1. Do you think those three criteria proper?

☐ Yes

☐ No

If no, please kindly give your recommendations on revising those criteria:

____

____

____

____

____

2. Are those criteria enough to select the top 5 routes?

☐ Yes

☐ No

If no, please kindly specify which criterion/criteria would you advice to add:

3. Please kindly specify pair wise importance of criteria:

(If you added some criteria, please skip this table and proceed to the next one on the next page)

	9:1	8:1	7:1	6:1	5:1	4:1	3:1	2:1	1:1	1:2	1:3	1:4	1:5	1:6	1:7	1:8	1:9	
RP																		*RL*
RP																		*RC*
RL																		*RC*

(If you didn't add any new criteria, please skip this table)

	9:1	8:1	7:1	6:1	5:1	4:1	3:1	2:1	1:1	1:2	1:3	1:4	1:5	1:6	1:7	1:8	1:9	

Thank you for spending your valuable time!

Appendix 2 AHP questionnaire #2

Dear Sir/Madam,

My name is Anton Hagen and I am doing my master thesis research on the topic “Planning a High Speed Rail Route in an Emerging Country: A Case Study of Ukraine”. In this study I’m going to develop the optimal *high speed rail* (HSR) route between two major cities in Ukraine. At the moment I have five alternative routes that where generated during the first stage of this study and now I have to make a decision about the optimal route, according to the following objectives:

- **Maximize** estimated travel demand (ETD) along the route, criterion:
  - estimated travel demand (TD)
- **Minimize** construction complexity (CC), criteria:
  - built-up area crossing index (BAI)
  - water body crossing index (WBI)
  - bad slope index (BSI)
- **Minimize** external effects (EE), criteria:
  - living area crossing index (LAI)
  - protected natural zone crossing index (PNZI)

All the five indexes of the second and third objectives are computed by tracing routes in GIS using the topographical map and digital elevation data as a basemap. Each route section is examined for crossing sensitive areas such as built-up areas, water bodies or protected natural zones and the length of each crossing is calculated. These indexes are calculated as follows:

$$I_{BA} = \frac{L_{RA} + 2L_{UA} + 2L_{IA}}{L}$$ Where I_{BA} is built-up area crossing index; L_{RA}, L_{UA}, L_{IA} are total lengths of rural, urban and industrial area crossings respectively; L is a route length.	$$I_{WB} = \frac{L_{WB}}{L}$$ Where I_{WB} is water body crossing index; L_{WB} is total length of water body crossings; L is a route length.
$$I_{LA} = \frac{L_{RA} + 2L_{UA}}{L}$$ Where I_{LA} is living area crossing index; L_{RA}, L_{UA} are total lengths of rural and urban area crossings respectively; L is a route length.	$$I_{PNZ} = \frac{L_{PNZ}}{L}$$ Where I_{PNZ} is protected natural zone crossing index; L_{PNZ} is total length of protected natural zone crossings; L is a route length.

Bad slope index is calculated as number of slopes that exceed 3.5% divided by route length:

$$I_{BS} = \frac{N(s_i > 0.035)}{L}$$

where I_{BS}: bad slope index;

s_i: slope value of the route section *i* that lies between vertexes *i-1* and *i* along the route;

L: route length;

$N(s_i > 0.035)$: number of route sections with slopes over the critical value.

Please help me with my thesis and answer the following questions.

4. Do you think those three objectives and their criteria proper?

☐ Yes

☐ No

If no, please kindly give your recommendations on revising those criteria:

____

____

____

____

____

5. Are those criteria enough to select the optimal route?

☐ Yes

☐ No

If no, please kindly specify which criterion/criteria would you advice to add:

____

____

____

____

____

____

____

____

6. Please kindly specify pair wise importance of the objectives:

- Estimated travel demand (ETD), **maximize**.
- Construction complexity (CC), **minimize**.
- External effects (EE), **minimize**.

(If you made changes that require another table, please skip this table and proceed to the next one)

	9:1	8:1	7:1	6:1	5:1	4:1	3:1	2:1	1:1	1:2	1:3	1:4	1:5	1:6	1:7	1:8	1:9	
ETD																		*CC*
ETD																		*EE*
CC																		*EE*

	9:1	8:1	7:1	6:1	5:1	4:1	3:1	2:1	1:1	1:2	1:3	1:4	1:5	1:6	1:7	1:8	1:9	

7. Please kindly specify pair wise importance of the criteria of **construction complexity**:

   a. build-up area crossing index (BAI)
   b. water body crossing index (WBI)
   c. bad slope index (BSI)

(If you made changes that require another table, please skip this table and proceed to the next one)

	9:1	8:1	7:1	6:1	5:1	4:1	3:1	2:1	1:1	1:2	1:3	1:4	1:5	1:6	1:7	1:8	1:9	
BAI																		*WBI*
BAI																		*BSI*
WBI																		*BSI*

	9:1	8:1	7:1	6:1	5:1	4:1	3:1	2:1	1:1	1:2	1:3	1:4	1:5	1:6	1:7	1:8	1:9	

8. Please kindly specify pair wise importance of the criteria of **external effects**:
   a. living area crossing index (LAI)
   b. protected natural zone crossing index (PNZI)

(If you made changes that require another table, please skip this table and proceed to the next one)

	9:1	8:1	7:1	6:1	5:1	4:1	3:1	2:1	1:1	1:2	1:3	1:4	1:5	1:6	1:7	1:8	1:9	
LAI																		*PNZI*

	9:1	8:1	7:1	6:1	5:1	4:1	3:1	2:1	1:1	1:2	1:3	1:4	1:5	1:6	1:7	1:8	1:9	

Thank you for spending your valuable time!

Appendix 3 Route demand data

Table 29 Route 1 demand

HSR Demand pass-km								
	Kyiv	Cherkasy	Kremenchuk	Dniprodzer-zhyns'k	Dnipropetrovs'k	Zaporizhzhya	Donetsk	**Total**
Kyiv	0	23 119 149	19 146 941	14 718 630	83 631 060	82 612 092	78 422 205	**301 650 076**
Cherkasy	86 513 184	0	6 089 494	4 829 806	27 629 305	27 587 416	26 544 310	**179 193 515**
Kremenchuk	37 846 688	3 216 616	0	4 633 986	26 983 803	10 840 201	27 023 657	**110 544 951**
Dniprodzer-zhyns'k	37 054 545	3 249 326	5 902 015	0	22 301 081	9 654 380	9 810 800	**87 972 147**
Dnipropet-rovs'k	92 579 091	8 173 439	15 111 932	3 861 159	0	22 969 005	23 952 439	**166 647 064**
Zaporizhzhya	81 518 644	7 274 680	5 411 556	3 784 109	20 474 353	0	19 915 262	**138 378 604**
Donetsk	111 219 207	10 060 087	19 389 043	5 526 766	30 686 345	28 622 887	0	**205 504 335**
Total	**446 731 358**	**55 093 297**	**71 050 981**	**37 354 456**	**211 705 948**	**182 285 979**	**185 668 672**	**1 189 890 691**

Table 30 Route 2 demand

HSR Demand pass-km									
	Kyiv	Cherkasy	Kirovohrad	Kryvy Rih	Dniprodzer-zhyns'k	Dnipropet-rovs'k	Zaporizh-zhya	Donetsk	**Total**
Kyiv	0	23 119 149	4 045 619	84 178 368	15 084 045	85 538 317	84 239 378	79 523 573	**375 728 448**
Cherkasy	86 513 184	0	501 089	27 554 373	5 024 288	28 601 324	28 359 026	27 019 682	**203 572 966**
Kirovohrad	28 530 126	944 329	0	20 367 180	1 507 990	8 629 194	21 909 158	21 074 463	**102 962 440**
Kryvy Rih	84 692 895	7 408 436	2 905 753	0	4 016 769	23 314 818	60 279 623	23 210 289	**205 828 582**
Dniprodzer-zhyns'k	37 974 485	3 380 167	538 337	10 050 888	0	22 301 081	9 675 374	9 810 800	**93 731 132**
Dnipropet-	94 690 413	8 460 986	1 354 555	25 652 569	3 861 159	0	23 040 620	23 952 439	**181 012 741**

rovs'k									
Zaporizhzhya	83 124 392	7 478 150	3 065 634	59 120 402	3 792 338	20 538 191	0	19 888 610	**197 007 716**
Donetsk	112 781 179	10 240 249	4 238 171	32 717 101	5 526 766	30 686 345	28 584 582	0	**224 774 393**
Total	**528 306 674**	**61 031 467**	**16 649 157**	**259 640 879**	**38 813 355**	**219 609 270**	**256 087 760**	**204 479 855**	**1 584 618 417**

Table 31 Route 3 demand

HSR Demand pass-km									
	Kyiv	Cherkasy	Kirovohrad	Kryvy Rih	Dniprodzer-zhyns'k	Dnipropet-rovs'k	Pavlohrad	Donetsk	**Total**
Kyiv	0	23 119 149	4 045 619	84 178 368	15 084 045	85 538 317	58 174 363	78 840 918	**348 980 778**
Cherkasy	86 513 184	0	501 089	27 554 373	5 024 288	28 601 324	7 699 152	26 726 828	**182 620 239**
Kirovohrad	28 530 126	944 329	0	20 367 180	1 507 990	8 629 194	5 937 336	20 800 231	**86 716 386**
Kryvy Rih	84 692 895	7 408 436	2 905 753	0	4 016 769	23 314 818	6 408 892	22 829 150	**151 576 712**
Dniprodzer-zhyns'k	37 974 485	3 380 167	538 337	10 050 888	0	22 301 081	6 562 642	9 589 614	**90 397 214**
Dnipropet-rovs'k	94 690 413	8 460 986	1 354 555	25 652 569	3 861 159	0	6 085 710	23 333 816	**163 439 208**
Pavlohrad	52 683 357	1 863 263	762 456	5 768 704	2 360 731	4 978 606	0	12 368 437	**80 785 553**
Donetsk	111 813 030	10 129 260	4 183 022	32 179 849	5 402 164	29 893 805	19 369 293	0	**212 970 424**
Total	**496 897 490**	**55 305 591**	**14 290 830**	**205 751 930**	**37 257 146**	**203 257 145**	**110 237 388**	**194 488 995**	**1 317 486 515**

Table 32 Route 4 demand

HSR Demand pass-km									
	Kyiv	Cherkasy	Kremenchuk	Poltava	Kharkiv	Slovyans'k	Kramantors'k	Donetsk	**Total**
Kyiv	0	23 119 149	19 146 941	9 432 646	78 556 584	1 401 303	4 828 964	79 166 147	**215 651 734**
Cherkasy	86 513 184	0	2 397 738	3 097 569	26 260 204	473 765	1 633 542	26 867 044	**147 243 046**
Kremenchuk	37 846 688	1 266 543	0	1 172 764	10 336 876	481 607	1 661 795	27 441 311	**80 207 584**

Poltava	39 034 177	3 425 485	2 455 239	0	9 696 154	183 380	633 426	10 517 636	**65 945 497**
Kharkiv	114 919 605	10 265 951	7 650 193	3 427 675	0	492 238	1 705 648	28 738 457	**167 199 765**
Slovyans'k	21 642 761	1 955 387	3 763 091	684 416	5 196 899	0	233 305	4 695 523	**38 171 382**
Kramantors'k	30 552 266	2 761 909	5 319 102	968 443	7 376 789	95 572	0	6 516 917	**53 590 998**
Donetsk	112 274 273	10 182 400	19 688 703	3 604 520	27 860 740	431 166	1 460 809	0	**175 502 612**
Total	**442 782 954**	**52 976 824**	**60 421 006**	**22 388 033**	**165 284 246**	**3 559 030**	**12 157 490**	**183 943 035**	**943 512 618**

Table 33 Route 5 demand

HSR Demand pass-km								
	Kyiv	Cherkasy	Kremenchuk	Dniprodzer-zhyns'k	Dnipropetrovs'k	Zaporizhzhya	Donetsk	**Total**
Kyiv	0	0	8 249 233	8 495 954	51 429 687	4 741 347	58 191 117	**131 107 337**
Cherkasy	152 001 027	0	0	2 152 037	14 534 204	579 522	19 568 930	**188 835 719**
Kremenchuk	50 951 131	5 801 706	0	0	7 445 030	185 898	19 329 878	**83 713 643**
Dniprodzerzhyns'k	40 685 956	3 801 000	8 150 912	0	0	517 490	8 221 623	**61 376 979**
Dnipropetrovs'k	106 758 512	10 209 094	22 266 740	10 486 870	0	0	16 671 482	**166 392 698**
Zaporizhzhya	68 815 843	2 714 433	2 420 737	6 093 068	17 436 814	0	0	**97 480 894**
Donetsk	0	0	0	0	0	0	0	**0**
Total	**419 212 469**	**22 526 234**	**41 087 621**	**27 227 927**	**90 845 735**	**6 024 256**	**121 983 030**	728 907 272