Acta logistica - International Scientific Journal about Logistics
Volume: 7 2020 Issue: 2 Pages: 85-93 ISSN 1339-5629

doi:10.22306/al.v7i2.161

Received: 25 Feb. 2020
Accepted: 02 Apr. 2020

ADJUSTING DIRECT DISTANCE TO ROAD FOR V4 COUNTRIES

Jakub Dyntar
Department of Business Administration and Management, Faculty of Economics, Technical University of Liberec, Voroněžská 13, 460 01 Liberec, Czech Republic, EU, jakub.dyntar@tul.cz
Dana Strachotová
Department of Economics and Management, The University of Chemistry and Technology Prague, Technická 5, 166 28 Praha 6, Czech Republic, EU, strachod@vscht.cz
Marek Botek
Department of Economics and Management, The University of Chemistry and Technology Prague, Technická 5, 166 28 Praha 6, Czech Republic, EU, botekm@vscht.cz (corresponding author)

Keywords: geographical information system, data mining, vehicle routing, distribution, supply chain management
Abstract: This article describes the determination of the direct distance correction factor that reflects the actual density of the road network in V4 countries. V4 countries are the Czech Republic (CZ), Slovakia (SK), Poland (PL) and Hungary (HUN). No correction factors reflecting road density among major population places were still available for V4 countries. Three-level administrative classification and data from statistical offices concerning population density on 31 December 2017 was used. In MS-Excel was designed functions for obtaining coordinates of selected places and road distances using queries to Web Map Service Google Maps (WMS). Road distances obtained by queries represent the fastest connection on the road. The great-circle distance and spherical triangle were used to calculate direct distances from coordinates. The places were selected using ABC analysis based on the population sample and it was reduced so that the monthly limit of queries to WMS was not exceeded. The obtained values of correction factors can be used in vehicle routing. For the smallest classification items, in CZ, for 273 places with a population of over 5 000, the average values of the correction factor (k_{avg}) range between 1.277 and 1.326. In SK, for 245 places with a population of over 3 000, k_{avg} ranges between 1.424 and 1.446. In PL, for 376 places with a population of over 20 000, k_{avg} ranges between 1.206 and 1.285. Finally, in HUN, for 287 places with a population of over 5 000, k_{avg} ranges between 1.301 and 1.345.

1 Introduction

The design of an efficient distribution system is an important part of the modern concept of material flow management commonly referred to as supply chain management [1]. Supply chain management (SCM) is defined as a material, information and money flow planning in a network of mutually linked organizations that add value to goods and services intending to satisfy the end consumer´s needs [2]. Seen from the perspective of processes, SCM includes planning, purchase, production, and distribution [3].

The main goal of the distribution is to supply customers with the required goods of the required quality and promptly. Distribution is a link between production and customers [4]. When designing distribution systems, companies typically decide about the number and location of warehouses [5] and the way how goods are transported to customers from warehouses (i.e. vehicle routing) so that distribution costs are as low as possible while maintaining the required service level [6]. The service level is represented particularly with the lead time [7]. The trend in the current market environment is extreme customers' pressure on lead time shortening [8]. It is important to model the whole supply chain to satisfied customers´ requirements [9].

A wide range of model approaches can be used to deal with the issue of vehicle routing. For instance, literature mentions the use of tabu search heuristic [10], simulated annealing algorithm [11] and genetic algorithm [12]. Another very efficient algorithm is the savings heuristic [13] and its modifications for vehicle routing with restrictions such as time windows [14] or vehicle capacity [15, 16]. All these algorithms are very influenced by the fact, that the size of the optimization model which results from its formulation grows extremely rapidly as the number of customers increases [17]. Also, it is necessary to distinguish, if the distance or time should be optimized [16].

What is common for the application of the above-stated models in vehicle routing is the necessity to take into account the distances among warehouses and customers. Warehouses and customers are most often represented with an address that indicates their location. Having an address means that geographic coordinates may be obtained from a geographic information system (GIS), typically using queries to a web map service (WMS). GIS is an organized aggregate of information technology, software and geographic data that allows for efficient obtaining, storing, updating, analyzing, transferring and viewing all kinds of geographic information [18]. WMS works on the client-server principle and it allows for sharing geographic information in the form of grid maps on the Internet [19].

If coordinates indicating the location of warehouses and customers are available, the direct distances among them can be calculated using, for example, the great-circle distance and spherical triangle [20]. However, such distances do not reflect the actual density of transportation routes in the regions where distribution takes place and they lead to distribution costs distortion. Actual distances can be obtained in two following ways:

- WMS direct query at the actual distance between two places linked by some kind of transport route,
- adjustment of direct distance using the correction factor that reflects the actual density of the transport network.

A drawback of the first way is a frequent limitation of WMS to the number of queries that can be made within a certain time interval as well as the fact that when queries exceed a certain number, the service is usually charged for. If distribution routes among *n* places using the savings algorithm is designed that works with a matrix of distances consisting of *n* lines and *n* columns, presuming that a distance between two places is the same as the distance on the way back, in order to fill the matrix of distances it would be necessary to make such an amount of queries to WMS that corresponds with the sum of an arithmetic line where the first term equals 1, the last term equals *n-1* and the difference is 1. If it is taken into consideration the fact that distribution commonly occurs between hundreds of places and vehicle routing is often performed daily, such an amount of queries to WMS presents a big burden. By contrast, when using the other option, it suffices to obtain only *n* geographic coordinates through queries to WMS.

The objective of this paper is to determine the correction factor of direct distance that reflects the actual density of the road transport network in V4 countries. V4 countries are the Czech Republic (CZ), Slovakia (SK), Poland (PL) and Hungary (HUN).

2 Materials and Methods

Administrative classification of V4 countries in 3 levels (see Table 1) together with data from statistical offices of those countries concerning numbers of population in L_3 on 31 December 2017 were used to determine the correction factor of direct distance (*k*), there was used.

Table 1 Administrative classification of V4 countries used for determining k

Level / Country	CZ	SK	PL	HUN
L_1	kraj	kraj	województwo	megye
L_2	okres	okres	powiat	kistérség
L_3	obec	obec	gmina	település

L3 components are first arranged in MS Excel according to the size of population and an ABC analysis is performed for each country following the share of the population living in *L3* components of the total population of each country. Group A includes *L3* components that contribute to the total number of each country´s population with 80%, Group B includes another 15% and Group C the remaining 5%.

After conducting an ABC analysis, *L3* components with the highest number of the population are involved in the calculation of k so that the monthly limit of queries to WMS Google Maps for three unique API keys is not exceeded. WMS Google Maps is employed for obtaining coordinates and route distances. *L3* components included in k calculation are shown in Table 2.

Table 2 L3 components included for determining k and the size of the population living in them

Country	No. of L_3 components	Population in L_3 components
CZ	273	>5 000
SK	245	>3 000
PL	376	>20 000
HUN	287	>5 000

The share of L_3 components included in *k* calculation in CZ is approx. 61%, in SK and PL it is around 62% and in HUN it is approx. 68% of the country´s population. In all monitored countries, the selection only included L_3 components from Group A. These are generally places where supply and demand are concentrated and with community facilities such as shops and restaurants are better developed. That means that such places are interesting from the perspective of distribution.

"Function1" is created in MS Excel (see Appendix 1) that uses Google Maps API Geocoding as a platform to calculate coordinates. The platform was accessed in 1/2019. The argument of the function (i.e. Address) used is the following string:

Table 3 Function1 - argument address

Country	L_1	L_2	L_3	Address
CZ	Středočeský kraj	Kladno	Slaný	L3&", okres "&L2&", "&L1
SK	Košický	Trebišov	Sečovce	L3&", okres "&L2&", "&L1&" kraj"
PL	Kujawsko-pomorskie		Tuchola	L3&", "&L1&" województwo"
HUN	Somogy	Fonyódi	Balatonlelle	L3&", "&L2&", "&L1

For instance, for the L_3 component from Table 3, situated in CZ, Function1 returns string 50.2304622, 14.0869439. After splitting the string into a part indicating latitude and longitude and once they are transformed to a number, L_3 coordinates are plotted on the graph and visual check of their location is performed in the administrative classification of L_1 (see Table 4).

Table 4 Administrative classification of L1 components in V4 countries

No./Country	CZ	SK	PL	HUN
1	Jihočeský	Banskobystrický	Dolnośląskie	Bács-Kiskun
2	Jihomoravský	Bratislavský	Kujawsko-pomorskie	Baranya
3	Karlovarský	Košický	Łódzkie	Békés
4	Královéhradecký	Nitranský	Lubelskie	Borsod-Abaúj-Zemplén
5	Liberecký	Prešovský	Lubuskie	Budapest
6	Moravskoslezský	Trenčínský	Małopolskie	Csongrád
7	Olomoucký	Trnavský	Mazowieckie	Fejér
8	Pardubický	Žilinský	Opolskie	Győr-Moson-Sopron
9	Plzeňský	x	Podkarpackie	Hajdú-Bihar
10	Praha	x	Podlaskie	Heves
11	Středočeský	x	Pomorskie	Jász-Nagykun-Szolnok
12	Ústecký	x	Śląskie	Komárom-Esztergom
13	Vysočina	x	Świętokrzyskie	Nógrád
14	Zlínský	x	Warmińsko-mazurskie	Pest
15	X	x	Wielkopolskie	Somogy
16	X	x	Zachodniopomorskie	Szabolcs-Szatmár-Bereg
17	X	x	X	Tolna
18	X	x	X	Vas
19	X	x	X	Veszprém
20	X	x	X	Zala

Direct distance D_1 between two L_3 components is always determined in MS Excel for two locations with a higher and lower number of the population using the following formula (1):

$$=6378,135*ARCCOS(COS(RADIANS(90-(x_2)))*COS(RADIANS(90-(x_1)))+SIN(RADIANS(90-(x_2)))*SIN(RADIANS(90-(x_1)))*COS(RADIANS((y_2-y_1)))), \quad (1)$$

where x_1 and y_1 are latitude and longitude of the location with higher number of population; x_2 and y_2 are latitude and longitude of the location with lower number of population.

In the case of CZ 37 128 direct distances are calculated, for SK it is 29 890, for PL 70 500, and for HUN 41 041.

“Function2” is created in MS Excel (see Appendix 2) that uses Google Maps API Directions as a platform to calculate the route distance D2 between two L3 components. Same as in the case of obtaining L_3 coordinates, the API Directions platform was accessed in 1/2019. The arguments of the function (i.e. Origin and Destination) used are strings in the form coming from Function1. Thanks to that, direct as well as route distances correspond with places of the same coordinates. Achieved D_2 represents the fastest route between two L_3 components on the road. That means that L_3 connection using motorways and main roads is preferred when queries are made to API Directions.

The direct distance correction factor when calculating road distance is defined for two L_3 components in the following way (2):

$$k = D2 / D1 \tag{2}$$

For each country, values of correction factors are arranged in a matrix, grouped according to the minimum number of population living in the L_3 components and averaged.

3 Results and Discussion

Table 5 shows the results of the ABC analysis by the share of the population living in L_3 components of the total population of each country.

Table 5 ABC analysis by the share of population living in L3 components of the total population of each country

		L_3 components in the group			
Group	**Group population/Country´s population**	**CZ**	**SK**	**PL**	**HUN**
A	80%	1 213	805	938	679
B	15%	2 474	1 075	899	1 142
C	5%	2 571	1 009	491	1 324
Total L_3 components in country		6 258	2 889	2 328	3 145

In CZ, 80% of the population lives in 1 213 L_3 components, which is approx. 19% of the total L_3 components in the administrative classification used. In SK, the same goes for 805 L_3 components with approx. 28%, in PL, it is 938 with approx. 40% and in HUN, it is 679 L_3 components with approx. 22% of the total number of L_3 components. In CZ, SK, and HUN the results of the ABC analysis are in line, to a great degree, with the Pareto principle (the 80/20 rule). In PL, the great number of municipalities classified in Group A is due to L_3 administrative classification when certain L_3 components unify populations of administrative centers and their related municipalities in the vicinity into one.

L_3 coordinates from Group A included in k calculation are plotted in graphs in Fig. 1 – 4.

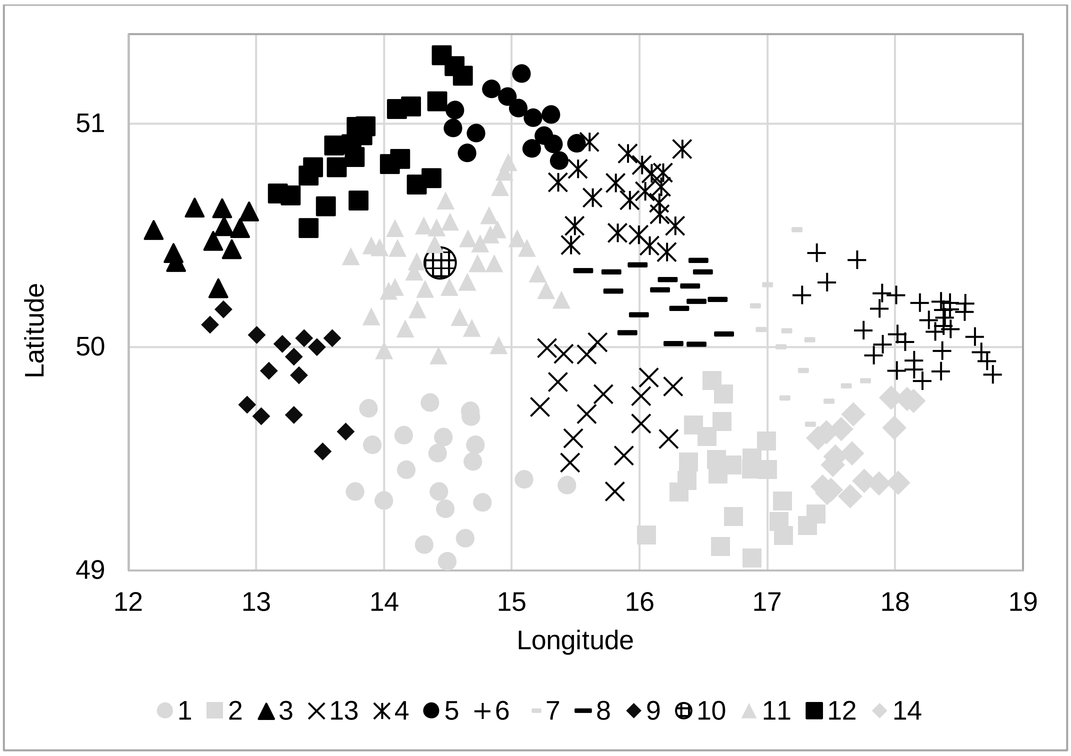

Figure 1 Location of L3 included in k calculation in L1 - CZ

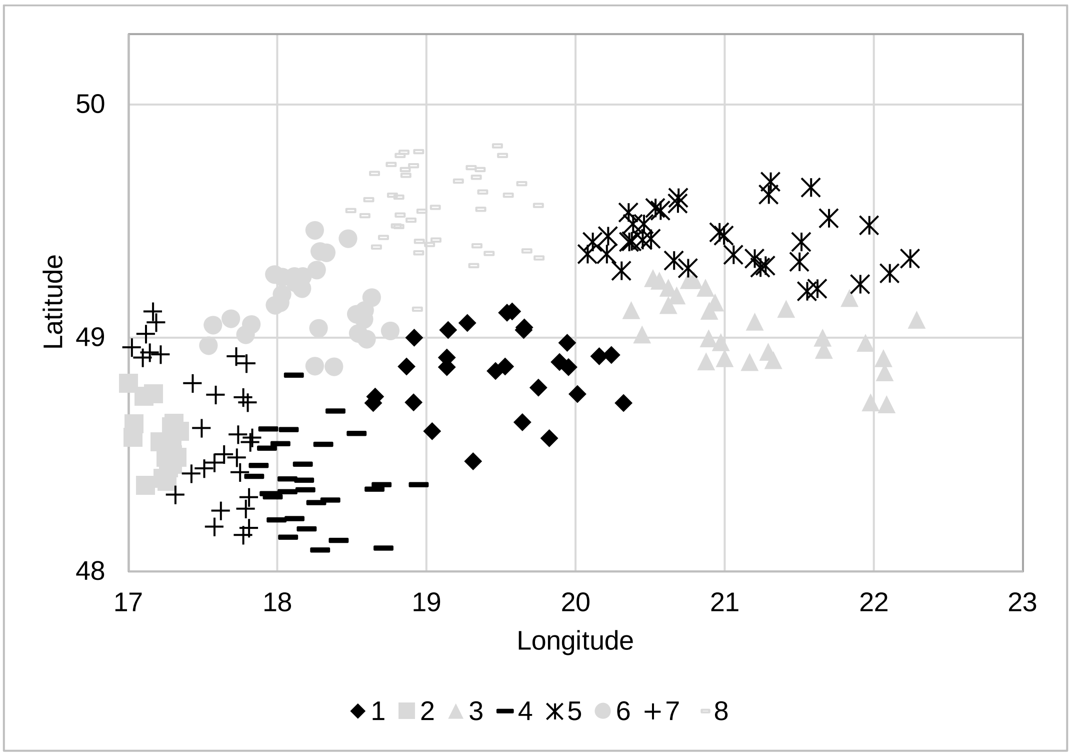

Figure 2 Location of L3 included in k calculation in L1 – SK

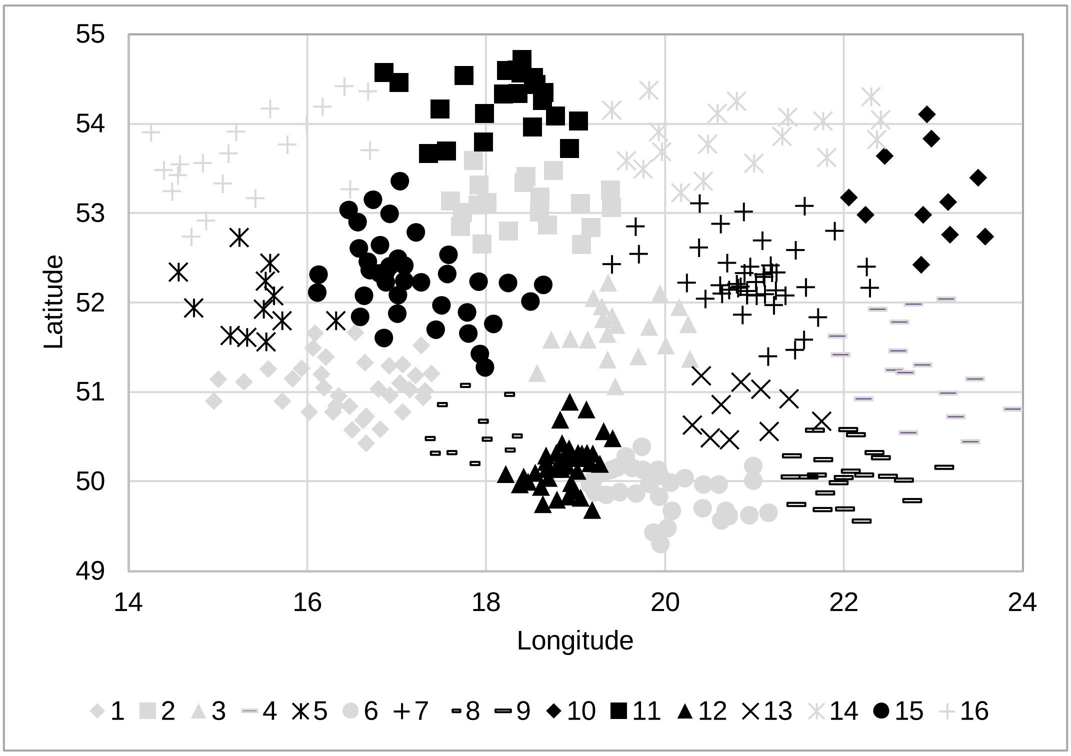

Figure 3 Location of L3 included in k calculation in L1 – PL

Acta logistica - International Scientific Journal about Logistics
Volume: 7 2020 Issue: 2 Pages: 85-93 ISSN 1339-5629

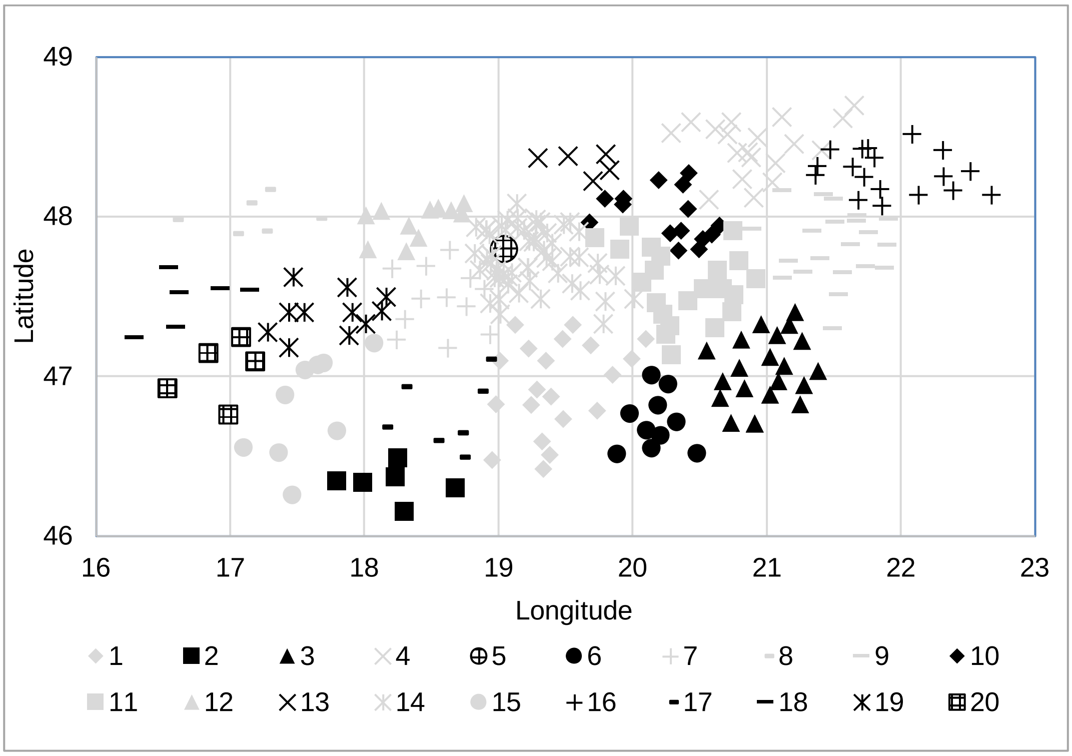

Figure 4 Location of L3 included in k calculation in L1 - HUN

It is important to note that L_3 components included in k calculation are only a part of Group A (see Table 2), which is due to the existence of a limited number of queries concerning coordinates and route distances of used WMS. In all the countries placement of L_3 within L_1 forms a compact area, while points mark L_3 locations, from two different L_1 components they only overlap at the frontier of L_1. Locations of L_1 components correspond with the actual locations of such administrative units in the real map and in the general view, they correspond with the actual shape of the countries.

Average values of correction factors for L_3 groups that differ with the minimum number of population living in L_3 components are shown in Table 6.

Table 6 Average k values for L3 groups differing with the minimum population living in L3 components

Country	CZ		SK		PL		HUN	
Population in L_3	No. of L_3	k_{avg}	No. of L_3	k_{avg}	No. of L_3	k_{avg}	No. of L_3	k_{avg}
>= 500 000	x	x	x	x	5	1.206	x	x
>= 250 000	x	x	x	x	11	1.231	x	x
>= 100 000	6	1.277	2	1.446	41	1.257	9	1.324
>= 50 000	18	1.322	10	1.428	111	1.275	21	1.301
>= 20 000	62	1.326	38	1.427	376	1.285	61	1.326
> =10 000	131	1.319	72	1.430	x	x	141	1.334
>= 7 000	190	1.323	97	1.425	x	x	202	1.343
>= 5 000	273	1.325	135	1.424	x	x	287	1.345
>= 3 000	x	x	245	1.426	x	x	x	x

In CZ, the average values of the correction factor (k_{avg}) range between 1.277 and 1.326. In SK, the k_{avg} values are between 1.424 and 1.446. In PL, the k_{avg} values are between 1.206 and 1.285, while in HUN the k_{avg} values they are between 1.301 and 1.345.

In CZ, PL, and HUN, the value grows together with the increasing number of L_3 components included in k_{avg} calculation (i.e. with the decreasing minimum number of population living in L_3). In SK, the trend is quite opposite, which is due to the mountainous landscape in the central part with only one backbone road connecting L_3 in the west and in the east that goes around the mountains and concentration of L_3 components with a higher number of population in the western part of the country.

The biggest difference between the route distance between L_3 with a higher and lower population as compared with direct distance (i.e. the difference given by the minimum and maximum k_{avg} value) is in PL (approx. 8%), followed by CZ (approx. 5%), HUN (approx. 4%), the difference in SK is very small (approx. 2%).

4 Conclusions

The objective of this article was to determine the correction factor of direct distance that reflects the actual density of the road network in V4 countries (the Czech Republic, Hungary, Poland, and Slovakia). The obtained values of correction factors correspond with the state of road networks in maps provided by WMS Google Maps in 1/2019. These values can be used for vehicle routing under the following assumptions:

1. Goods are distributed by vehicles in the road network.
2. Customers demanding goods are at locations with a population of more than 3 000 (in SK), 5 000 (in CZ and HUN) and 20 000 (in PL).
3. Locations, where demand is concentrated, are scattered all over the country. Therefore, goods are not distributed, for instance, exclusively (or predominantly) in the capital city or in a relatively small region where actual distances may be distorted by the presence of natural obstacles such as mountains or great rivers.
4. When designing the distribution route, the fastest connection is preferred, which means that motorways and/or main roads without the risk of free passage restriction for vehicles exceeding certain weight are primarily used.

Obtained correction factors can be primarily used on a strategic level to support decisions about distribution system (re)design based on historical demand data. By reducing the time necessary to calculate distances between the places with a demand it is possible to assess many different varieties of a distribution system structure design (i.e. appropriate number and a location of warehouses and a strategy of supplying the customers) in a short time. On the other hand, the usage of correction factors as a part of vehicle routing for operative planning can lead to the estimation of distances that is too optimistic. That is because of not considering real-time traffic data about rush hours or road closures or including the places of demand with a very small population.

References

[1] MENTZER, J.T., DEWITT, W., KEEBLER, J.S., MIN, S., NIX, N.W., SMITH, C.D., ZACHARIA, Z.G.: Defining Supply Chain Management, *Journal of Business Logistics*, Vol. 22, No. 2, pp. 1-25, 2001.

[2] STOCK, J.R., BOYER, S.L.: Developing a consensus definition of supply chain management: a qualitative study, *International Journal of Physical Distribution & Logistics Management*, Vol. 39, No. 8, pp. 690-711, 2009.

[3] COOPER, M.C., LAMBERT, D.M., PAGH, J.D.: Supply Chain Management: More Than a New Name for Logistics, *The International Journal of Logistics Management*, Vol. 8, No. 1, pp. 1-14, 1997.

[4] CHOPRA, S.: Designing the distribution network in a supply chain, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 39, No. 2, pp. 123-140, 2003. doi:10.1016/S1366-5545(02)00044-3

[5] STRAKA, M., BINDZAR, P., KADUKOVA, A.: Utilization of the multicriteria decision-making methods for the needs of mining industry, *Acta Montanistica Slovaca*, Vol. 19, No. 4, pp. 199-206, 2014.

[6] BRUNAUD, B., BASSETT, M.H., AGARWAL, A., WASSICK, J. M., GROSSMANN, I. E.: Efficient formulations for dynamic warehouse location under discrete transportation costs, *Computers & Chemical Engineering*, Vol. 111, No. March, pp. 311-323, 2018. doi:10.1016/j.compchemeng.2017.05.011

[7] ESKIGUN, E., UZSOY, R., PRECKEL, P.V., BEAUJON, G., KRISHNAN, S., TEW, J.D.: Outbound supply chain network design with mode selection, lead times and capacitated vehicle distribution centers, *European Journal of Operational Research*, Vol. 16, No. 1, pp. 182-206, 2005. doi:10.1016/j.ejor.2003.11.029

[8] de TREVILLE, S., SHAPIRO, R.D., HAMERI, A.P.: From supply chain to demand chain: the role of lead time reduction in improving demand chain performance, *Journal of Operations Management*, Vol. 2, No. 6, pp. 613-627, 2004. doi:10.1016/j.jom.2003.10.001

[9] GRAKOVA, E., LENORT, R., WICHER, P.: *Conceptual Framework for Computer Simulation of Supply Chain Resilience Building*, Infront Outback – Conference Proceedings CLC 2013: Carpathian Logistics Congress in Cracow, Poland, pp. 429-433, *2013*.

[10] HO, S.C., HAUGLAND, D.: A tabu search heuristic for the vehicle routing problem with time windows and split deliveries, *Computers & Operations Research*, Vol. 31 No. 12, pp. 1947-1964, 2004. doi:10.1016/S0305-0548(03)00155-2

[11] OSMAN, I.H.: Metastrategy simulated annealing and tabu search algorithms for the vehicle routing

problem, *Annals of Operations Research*, Vol. 41, No. 4, pp. 421-451, 1993. doi:10.1007/BF02023004

[12] BAKER, B.M., AYECHEW, M.A.: A genetic algorithm for the vehicle routing problem, *Computers & Operations Research*, Vol. 3 No. 5, pp. 787-800, 2003. doi:10.1016/S0305-0548(02)00051-5

[13] TENAHUA, A., OLIVARES-BENITEZ, E., SÁNCHEZ-PARIDA, D., CABALLERO-MORALES, S.O.: ILS Metaheuristic to solve the Periodic Vehicle Routing Problem, *International Journal of Combinatorial Optimization Problems and Informatics*, Vol. 9, No. 3, pp. 55-63, 2018.

[14] LI, H.Q., ZHANG, L., LV, T., CHANG, X.Y.: The two-echelon time-constrained vehicle routing problem in linehaul-delivery systems, *Transportation Research Part B-Methodological*, Vol. 94, No. December, pp. 169-188, 2016. doi:10.1016/j.trb.2016.09.012

[15] VIGO, D.: A heuristic algorithm for the asymmetric capacitated vehicle routing problem, *European Journal of Operational Research*, Vol. 89, No. 1, pp. 108-126, 1996. doi:10.1016/S0377-2217(96)90060-0.

[16] ZHAO, T. L., YANG, Y. J., WANG, E.: Minimizing the average arriving distance in carpooling, *International Journal of Distributed Sensor Networks*, Vol. 16, No. 1, 2020. doi:10.1177/1550147719899369

[17] WANG, J. L., JAGANNATHAN, A. K. R., ZUO, X., MURRAY, C. C. Two-layer simulated annealing and tabu search heuristics for a vehicle routing problem with cross docks and split deliveries. *Computers & Industrial Engineering*, Vol. 112, No. October, pp. 84-98, 2017. doi:10.1016/j.cie.2017.07.031

[18] KLATZKY, R.L., LOOMIS, J.M., SPEIGLE, J.O.N., TIETZ, J.: A geographical information system for a GPS based personal guidance system, *International Journal of Geographical Information Science*, Vol. 12, No. 7, pp. 727-749, 1998. doi: 10.1080/136588198241635

[19] BATTY, M., HUDSON-SMITH, A., MILTON, R. AND CROOKS, A.: Map mashups, Web 2.0 and the GIS revolution, *Annals of GIS*, Vol. 16, No. 1, pp. 1-13, 2010.

[20] LEE, J.E., SUNG, J.H., WARD, W.B., FOS, P.J., LEE, W.J., KIM, J.C.: Utilization of the emergency room: impact of geographic distance, *Geospatial Health*, Vol. 1, No. 2, pp. 243-253, 2007. doi:10.4081/gh.2007.272

Review process
Single-blind peer review process.

Appendix 1

```
Function Function1(Address As String) As String
 Dim Request As New XMLHTTP60
 Dim Results As New DOMDocument60
 Dim StatusNode As IXMLDOMNode
 Dim LatitudeNode As IXMLDOMNode
 Dim LongitudeNode As IXMLDOMNode
  On Error GoTo exit
    Request.Open "GET", "https://maps.googleapis.com/maps/api/geocode/xml?" _
    & "&address=" & Application.EncodeURL(Address) & "&key=A*zaS*DQ**y***lPX3u****y7ogt*J1**u*k9*",
False
    Request.send
    Results.LoadXML Request.responseText
    Set LatitudeNode = Results.SelectSingleNode("//result/geometry/location/lat")
    Set LongitudeNode = Results.SelectSingleNode("//result/geometry/location/lng")
    Function1 = LatitudeNode.Text & ", " & LongitudeNode.Text
exit:
 Set StatusNode = Nothing
 Set LatitudeNode = Nothing
 Set LongitudeNode = Nothing
 Set Results = Nothing
 Set Request = Nothing
End Function
```

Appendix 2

```
Function Function2(Origin As String, Destination As String) As Double
 Dim myRequest As XMLHTTP60
```

```
 Dim myDomDoc As DOMDocument60
 Dim distanceNode As IXMLDOMNode
  On Error GoTo exit
   Origin = WorksheetFunction.EncodeURL(Origin)
   Destination = WorksheetFunction.EncodeURL(Destination)
   Set myRequest = New XMLHTTP60
    myRequest.Open "GET", "https://maps.googleapis.com/maps/api/directions/xml?" _
    & "&origin=" & Origin & "&destination=" & Destination & "&sensor=false" &
"&key=A*zaS*DQ**y***lPX3u****y7ogt*J1**u*k9*", False
    myRequest.send
   Set myDomDoc = New DOMDocument60
    myDomDoc.LoadXML myRequest.responseText
   Set distanceNode = myDomDoc.SelectSingleNode("//leg/distance/value")
    If Not distanceNode Is Nothing Then Function2 = distanceNode.Text / 1000
exit:
 Set distanceNode = Nothing
 Set myDomDoc = Nothing
 Set myRequest = Nothing
End Function
```