Comparable Measures of Accessibility to Public Transport Using the General Transit Feed Specification

Jinjoo Bok 1,* and Youngsang Kwon 2,*

1 Department of Civil & Environmental Engineering, Seoul National University, Gwanak-ro 1, Gwanak-gu, Seoul 08826, Korea
2 Department of Civil and Environmental Engineering, Integrated Research Institute of Construction and Environmental Engineering, Seoul National University, Gwanak-ro 1, Gwanak-gu, Seoul 08826, Korea
* Correspondence: manjjew@gmail.com (J.B.); yskwon@snu.ac.kr (Y.K.); Tel.: +82-2-880-8200 (J.B. & Y.K.)

Abstract: Public transport plays a critical role in the sustainability of urban settings. The mass mobility and quality of urban lives can be improved by establishing public transport networks that are accessible to pedestrians within a reasonable walking distance. Accessibility to public transport is characterized by the ease with which inhabitants can reach means of transportation such as buses or metros. By measuring the degree of accessibility to public transport networks using a common data format, a comparative study can be conducted between different cities or metropolitan areas with different public transit systems. The General Transit Feed Specification (GTFS) by Google Developers allows this by offering a common format based on text files and sharing the data set voluntarily produced and contributed by the public transit agencies of many participating cities around the world. This paper suggests a method to assess and compare public transit accessibility in different urban areas using the GTFS feed and demographic data. To demonstrate the value of the new method, six examples of metropolitan areas and their public transit accessibility are presented and compared.

Keywords: accessibility; public transport; general transit feed specification; comparable measures; geographic information system

1. Introduction

Public transport in urban areas has gained greater attention in recent years for improving sustainability and the quality of urban life. The economic and environmental performance of cities can be enhanced by connecting resources to destinations effectively and facilitating mass mobility. Public transport networks are an important part of urban structure that enables other elements of urban interactions and circulations to be developed. Thus, monitoring and assessing public transit can be crucial for understanding an urban system and its major functions.

Public transport is one component of the modal split that includes walking, cycling, and private vehicles [1]. In contrast to private transport modes, public transit is considered to be open to a wider user group and is operated by public authorities or agents. Besides mass mobility, unique characteristics such as routes and schedules, locations of users, and the times of day when trips are made distinguish public transit from other modes of transportation and influence accessibility [2]. To measure accessibility, this study divides public transport into two broader groups: rail-based transit (metros or trains) and road-based transit (buses or trams).

The concept of accessibility can generally be defined as the ease with which inhabitants can reach their destinations [3]. However, similar but separate notions have been used in the last 40 years of
transportation research [4,5]. Terminological differences such as “access,” “accessibility,” “availability,” and “proximity” have been used in various studies with similar meaning [6–8]. A number of measures of accessibility have been proposed [9–14], but the measurements vary with the researchers’ perspectives and objectives [2,15,16]. While traditional studies on transportation accessibility have focused on transit as a means of reaching other urban services and areas of activities, recent studies define public transport as a key urban service to be accessed [17]. Therefore, the expression of accessibility to public transport rather than accessibility of public transport should be noted.

This study focuses on the temporal dimension of public transport as well as spatial aspect. The spatial factor of accessibility is identified as the specific distance from each node of transit. Physical distance to a node on a route is the most basic component of accessibility to public transport, as public transport needs to be physically reachable for a large number of inhabitants [18]. However, the proximity of users to transit stops is not sufficient to define accessibility; when and how frequently vehicles can be accessed are also decisive factors [6]. In terms of the modal split of public transport, both the peak service volume and average service volume per day have generally been used. Operational frequency has usually been measured as trips per hour by taking the average of scheduled headways along with the routes [19].

The most recent analytical studies focus on taking into account “time” in the geographic information science, transportation, and geography fields. With the innovation and implementation of information and transit technologies, the conventional geographic concepts and measuring methodologies have become insufficient in representing the real-time environment [20]. Integration of time and space by capturing individual mobility in everyday life has changed aggregated place-based perspectives into people-based perspectives, owing to the use of information and communication technology (ICT), location-aware technology (LAT), global positioning system (GPS), radiofrequency identification (RFID), and location-based services (LBS) [21–23].

On the other hand, other recent work from the perspective of large-scale public transit applications develops the integrated measuring concept of “connectivity” for assessing performance of multimodal and multilevel transit networks [24–26]. Associated attributes including nodes, lines, transfers, schedules, and spatial activity patterns are quantified as an indicator for decision-making and the allocation of funding in the most efficient way.

This study aims to explore an internationally comparable methodology for measuring accessibility to public transport with interoperable data sources and compatible spatial units for analysis that can be applied globally. The increasing requirements for global collaboration towards more sustainable urban structures and environments can be met by establishing common parameters for analysis among various cities and metropolitan areas with different dimensions and contexts. Achieving sustainable urban form and improving regional well-being are key challenges for current urban and regional policies, and public transport plays a vital role in addressing these issues.

In Section 2, we highlight the importance of a new transit data source and illustrate a framework to measure public transport accessibility. In Section 3, application of the method to one example of an Asian metropolitan area and five North American areas is presented. Finally, we conclude with a short summary and insights for future works.

2. Data and Framework

Many research efforts explore the importance of precise estimation of population related to public transit and enhance analysis methods. Choice of representation of proper transit elements and levels of population data aggregation that overlap considerably tend to alter the final results [27–29]. Although effects of scale and spatial representation issues can be addressed with new methods [28,29], as long as researchers apply census data (i.e., residential population distribution, not actual daytime population), there is the limitation of misrepresentation of actual population and over- or under-estimation of demand.
General Transit Feed Specification (GTFS) has been used in recent transit research since Google structured and launched the open platform in 2008. Despite the limitation of spatial coverage of worldwide GTFS data, especially in the Global South [30], these standardized transit data have great potential as a source for efficient spatiotemporal transit analyses by using scripts and database queries [31], especially for comparing research between various cities. Shortest path routes along the 2016 Sustainability the EU (Eurostat and European Commission Directorates General (EC-DG) Regio) (Figure 1) [34,35].

Meeting this demand involves accounting for international comparability in OECD works. Demand is growing for a common base to understand and assess urban areas to ensure population vary considerably across the FUAs.

The calculation of the public transport accessibility indicator for metropolitan areas consists of four steps: (1) data collection; (2) identification of service areas; (3) classification of accessibility levels; and (4) calculation of public transit coverage in terms of population. In addition, a common spatial unit of analysis is needed for different dimensional metropolitan areas before providing a methodology for measuring accessibility. This study applies Functional Urban Areas (FUAs), which were defined in 2012 by the Organization for Economic Cooperation and Development (OECD) in collaboration with the European Union (EU) for the purpose of international comparison.

2.1. Spatial Unit of Analysis—Functional Urban Areas (FUAs)

Demand is growing for a common base to understand and assess urban areas to ensure international comparability in OECD works. Meeting this demand involves accounting for administrative boundaries, the continuity of built-up areas, functional measures such as commuting rates, and the size of components to be aggregated. Recently, 1179 FUAs from 29 OECD countries were defined as urban spatial units including urban cores and hinterlands by the OECD in cooperation with the EU (Eurostat and European Commission Directorates General (EC-DG) Regio) (Figure 1) [34,35]. Figure 2 and Table 1 show six cases of FUA maps and profiles with public transit [36,37]. Both the size of the overall metropolitan area and population vary considerably across the FUAs.

![Figure 1. FUAs with population. (Source: OECD Metropolitan Explorer).](image_url)
Figure 2. Case FUAs’ maps and routes.
Table 1. Profiles of case FUAs.

<table>
<thead>
<tr>
<th>FUA</th>
<th>Chicago (US)</th>
<th>Portland (US)</th>
<th>Washington D.C. (US)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area (2010)/(km²)</td>
<td>17,465</td>
<td>17,657</td>
<td>12,085</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FUA</th>
<th>Vancouver (Canada)</th>
<th>Toronto (Canada)</th>
<th>Daejeon (Korea)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population (2010)</td>
<td>2,312,497</td>
<td>6,418,623</td>
<td>1,565,511</td>
</tr>
<tr>
<td>Area (2010)/(km²)</td>
<td>5064</td>
<td>14,413</td>
<td>996</td>
</tr>
</tbody>
</table>

Source: US Census, Daejeon Metropolitan City.

2.2. Data Collection—General Transit Feed Specification (GTFS)

The General Transit Feed Specification (GTFS) by Google Developers facilitates the publishing and sharing of public transit information of various cities and metropolitan areas by defining a common format. GTFS consists of a series of text files containing specific aspects of public transport data: stops, routes, trips, and other schedule information (Table 2) [38]. The GTFS feed is voluntarily produced and updated on the Internet by the public transport agencies of cities and municipalities around the world. Since first being launched in 2008, over 900 transit agencies have shared their public transportation feed in this interoperable way. While the major contributors are North American cities, there are growing numbers of European and Asian cities participating in the program (Table 3) [39].

Table 2. GTFS files and their contents.

<table>
<thead>
<tr>
<th>Filename</th>
<th>Defines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agency.txt</td>
<td>One or more transit agencies that provide the data in this feed.</td>
</tr>
<tr>
<td>Stopd.txt</td>
<td>Individual locations where vehicles pick up or drop off passengers.</td>
</tr>
<tr>
<td>Routes.txt</td>
<td>Transit routes. A route is a group of trips that are displayed to riders as a single service.</td>
</tr>
<tr>
<td>Trips.txt</td>
<td>Trips for each route. A trip is a sequence of two or more stops that occurs at a specific time.</td>
</tr>
<tr>
<td>Stop_times.txt</td>
<td>Times when a vehicle arrives at and departs from individual stops for each trip.</td>
</tr>
<tr>
<td>Calendar.txt</td>
<td>Dates for service IDs using a weekly schedule. Specifies when service starts and ends, as well as days of the week when service is available.</td>
</tr>
<tr>
<td>Calendar_dates.txt</td>
<td>Exceptions for the service IDs defined in calendar.txt. If calendar_dates.txt includes all dates of service, this file may be specified instead of calendar.txt.</td>
</tr>
<tr>
<td>Fare_attributes.txt</td>
<td>Fare information for a transit organization’s routes.</td>
</tr>
<tr>
<td>Fare_rules.txt</td>
<td>Rules for applying fare information for a transit organization’s routes.</td>
</tr>
<tr>
<td>Shapes.txt</td>
<td>Rules for drawing lines on a map to represent a transit organization’s routes.</td>
</tr>
<tr>
<td>Frequencies.txt</td>
<td>Headway (time between trips) for routes with variable frequency of service.</td>
</tr>
<tr>
<td>Transfers.txt</td>
<td>Rules for making connections at transfer points between routes.</td>
</tr>
<tr>
<td>Feed_info.txt</td>
<td>Additional information about the feed itself, including publisher, version, and expiration information.</td>
</tr>
</tbody>
</table>

Source: Google Static Transit.
In this study, transit data from Chicago, Portland, Washington, Vancouver, and Toronto were obtained through the GTFS platform. Daejeon had not yet been included in the list, so data for these FUAs were retrieved through an additional search. Although GTFS feeds do not necessarily contain all public transport data for certain cities or metropolitan areas, efficient availability of data with a common format enables comparative analysis of a number of cases.

2.3. Identification of Areas Serviced by Public Transit

A database file, where six standard GTFS files are imported and dynamically linked together with queries using Microsoft ACCESS 2010, has been created in order to generate an output table that contains, for every stop of every transit mode and for one specific date (8 October 2013), the frequency of departures for each hour.

All stops then have been imported as point features containing the average number of departures per hour during working-day hours (from 6 a.m. to 8 p.m.) along the pedestrian network. The stops located within 50 m from another stop were considered as one single cluster of stops with an aggregated hourly number of departures and new geographical mean center coordinates to avoid underestimating the effects of nearby stops on frequency.

Areas serviced by public transport are created by calculating pedestrian walking time through the Network Analyst Tools of ArcGIS 9.3 using the defined thresholds of 5 min of walking time for accessing road-based transit and 10 min of walking time for accessing rail-based transit (Figure 3). The walking speed applied at this stage is 1.1 m/s (almost 4 km/h), resulting in a 330 m network distance to road-based transit nodes and a 660 m network distance to rail-based transit nodes.

**Table 3. Participating agencies by continent.**

<table>
<thead>
<tr>
<th>Continent</th>
<th>North America</th>
<th>Europe</th>
<th>Asia</th>
<th>Central/South America</th>
<th>Africa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of participating agencies (%)</td>
<td>56.40</td>
<td>33.48</td>
<td>6.72</td>
<td>2.45</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Source: GTFS Data exchange.
2.4. Classification of Serviced Area Accessibility

Frequency was added as a temporal component for combined assessment after defining the serviced areas of public transport. Considering the road-based and rail-based modes, there are two steps to finalize mode-integrated classification for each metropolitan area: (1) classifying each serviced area for each mode by frequency data and (2) reclassification by aggregating mode-specific classes.

Frequency levels are determined by using the average amount of public transport service headways per hour. The mode-specific frequency classes have been defined as follows and applied for both modes (Table 4). Aggregated accessibility levels of public transport are divided into five classes (Table 5). Only the combination featuring high rail-based and road-based public transport results in a very high level of accessibility, while other combinations result in high, medium, low, and no access levels of accessibility.

Table 4. Classification of service areas by frequency.

<table>
<thead>
<tr>
<th>Level</th>
<th>Average Headway (from 6 a.m. to 8 p.m.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Less than 6 min</td>
</tr>
<tr>
<td>Medium</td>
<td>From 6 min to 15 min</td>
</tr>
<tr>
<td>Low</td>
<td>More than 15 min</td>
</tr>
<tr>
<td>No service</td>
<td>No operation</td>
</tr>
</tbody>
</table>

Table 5. Aggregated reclassification of service areas by frequency.

<table>
<thead>
<tr>
<th>Level of Rail-based Transit Service Area</th>
<th>High</th>
<th>Medium</th>
<th>Low</th>
<th>No Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of road-based transit service area</td>
<td>High</td>
<td>Very High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>No service</td>
<td>High</td>
<td>Medium</td>
<td>No service</td>
</tr>
</tbody>
</table>

2.5. Calculating the Population Share in Transit Service Areas

Ambient population data by Oak Ridge National Laboratory are measured for approximately one square kilometer, depending on the location’s distance from the Equator. The expected population, being a 24 h estimate, is an average for the day, regardless of the time of year, since calculations incorporate both diurnal and seasonal population movements. LandScan 2009 was used for the estimation of population-accessible transit in this study. Population in the catchment areas of each accessibility level can be calculated by overlapping the service areas layer and ambient population layers on the assumption of evenly distributed ambient population in a cell.

Comparison across metropolitan areas can be performed by calculating the shares of population living in catchment areas of each accessibility level. Six FUAs were selected as case areas: Daejeon (Korea); Chicago, Portland, and Washington (U.S.); and Toronto and Vancouver (Canada). The serviced population shares in both FUA areas and urban core areas were calculated according to the different service levels (ranging from very high to no service).

3. Results and Discussion

Table 6 shows the number of transit stops of all modes obtained by GTFS data as of November 2013 as well as populations and areas of six different FUAs. As it was introduced earlier, FUAs consist of both urban core and hinterlands, which are identified as major job provider/functional core areas and labor supply/residential areas, respectively. Buses are found to be the most highly accessible transit mode due to the fact that bus stops heavily outnumber those of any other transit modes in every FUA. The Portland FUA has the lowest density as well as the lowest number of people per each
bus stop among all FUAs, while the Chicago FUA shows the highest density of the North American FUAs, except the Daejeon FUA has the highest density of all.

Table 6. Transit profiles in FUAs.

<table>
<thead>
<tr>
<th>FUA</th>
<th>Population (2010)</th>
<th>Area (2010)/(km²)</th>
<th>Density (person/km²)</th>
<th>Stops or Stations</th>
<th>Persons per bus stop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicago (US)</td>
<td>9,315,366</td>
<td>17,465</td>
<td>533.4</td>
<td>Bus 11,238, Metro 135</td>
<td>828.9</td>
</tr>
<tr>
<td>Portland (US)</td>
<td>2,226,009</td>
<td>17,657</td>
<td>126.1</td>
<td>Bus 7550, Metro 4</td>
<td>294.8</td>
</tr>
<tr>
<td>Washington D.C. (US)</td>
<td>5,254,951</td>
<td>12,085</td>
<td>434.8</td>
<td>Bus 11,081, Metro 86</td>
<td>474.2</td>
</tr>
<tr>
<td>Vancouver (Canada)</td>
<td>2,312,497</td>
<td>5064</td>
<td>456.7</td>
<td>Bus 1007, Metro 135</td>
<td>2296.4</td>
</tr>
<tr>
<td>Toronto (Canada)</td>
<td>6,418,623</td>
<td>14,413</td>
<td>445.3</td>
<td>Metro 666, Metro 66</td>
<td>748.9</td>
</tr>
<tr>
<td>Daejeon (Korea)</td>
<td>1,565,511</td>
<td>996</td>
<td>1571.8</td>
<td>Metro 2027, Train 11</td>
<td>777.2</td>
</tr>
</tbody>
</table>

Figure 4 provides the geographic distribution of public transport catchment areas according to the accessibility levels in urban cores of the six FUAs. As few public transit stops and stations in the hinterlands of each FUA were observed, only those in the urban core are displayed. Even within urban cores, bus and tram stops are found to be more closely concentrated in the central areas of the urban cores. On the other hand, commuting train and metro stations are sparsely distributed in linear patterns towards the outskirts. These differences of spatial distribution according to transit modes are reflected on access time taken by riders on foot: 5 min for road-based transit and 10 min for rail-based transit.

Service areas, regardless of their levels of accessibility, refer to the areas where the estimated walking time for accessing public transport is no longer than 5 min for road-based transit and 10 min for rail-based transit. The grey areas, on the other hand, refer to areas where the estimated walking time for accessing public transport is longer than 5 min and 10 min, respectively. The service frequency of public transport was classified into the categories of very high, high, medium, and low. Among the four accessibility levels, the very high level service areas (navy areas) refer to the areas where a passenger who just missed a bus or train can take the next one of either transit modes within the next 6 min. Secondly, the high level service areas (dark green areas) represent the areas where the rider can take only one of the two transit modes within the next 6 min. Thirdly, the areas with the medium level of accessibility (light green areas) show service areas where the rider can take at least one of the two transit modes in 6 to 15 min. Lastly, the low level service areas (yellow areas) refer to the areas where the rider needs to wait for longer than 15 min for both transit modes, or where the rider cannot reach any nodes of either transit modes within 5 min and 10 min, respectively.

All six results show the highest levels of access are concentrated around central nodes in urban core areas, with lower levels radiating towards the margins of the urban cores and spreading sparsely into the hinterlands. Serviced areas are mostly located within the urban core areas. Despite considerable variation, a certain distribution pattern of larger populations and higher levels in urban cores is observed in all assessed areas. In comparison to Asian metropolitan areas, North American cities illustrate more intensively concentrated catchment areas in a smaller portion of areas within urban cores, which seems to correspond with the relatively extensive land use pattern and high dependency on automobiles.
The population share in serviced areas more clearly reveals the accessibility to public transport across the case regions (Figure 5). Daejeon has the largest share of population with high accessibility among the six FUAs. A relatively smaller area of the Daejeon FUA compared to the North American FUAs corresponds with a higher share of area of public transit catchments, with 22% of the whole Daejeon FUA area and 30% of the urban core area (Table 1 and Figure 6). Nevertheless, excessive differences between shares of population and area of each FUA reveal a centralized urban structure and transit networks of metropolitan areas. Except for Portland among the five North American FUAs, others show that around 70% of the population has no access to public transport in the metropolitan areas.
Figure 5. Share of population with access to public transport in FUAs.

Figure 6. Share of area with access to public transport in FUAs.

4. Conclusions

Ensuring public transport accessibility is an important task for sustainable urban development. This study focused on conceptualizing an integrated indicator and providing a viable methodology and data sources for measuring and comparing the accessibility to public transport for metropolitan areas. The proposed method applies spatial and temporal dimensions of accessibility based on a common spatial unit of analysis for comparison of the assessment results.

The methodology in this study differs from that of other studies in that it integrates several transit modes, including both on roads and on rails rather than only bus, and temporally classifies the service areas with actual headways of each mode and stop on a specific date. For the estimation of ridership, it introduces a novel concept and data, that of ambient population instead of census data in order to represent the actual daytime population distribution.

The results indicate how many people living in FUAs have relatively easy access to public transport in terms of both physical access and the level of service frequency provided. In comparison with another 14 European metropolitan area cases [40], Figure 7 shows that a larger share of the population in urban core areas of European metropolitan areas has access to public transit compared to North American areas; above 70% of the population in the European cities has some access to public transport. There was a striking difference between European and non-European cities, especially cities in North America. The general pattern shows larger population shares being serviced in urban cores, which also have higher levels of service frequency compared to the FUAs overall. These trends confirm the expectations in public transport development and land use patterns that more densely inhabited urban areas would tend to feature better public transport than the hinterlands.

Given that the major purpose of a comparative study is to provide insights for policy implementation and decision-making, the behavioral insights from this work are: the sustainable expansion of metropolitan areas in terms of efficient supply and management of public transport.
Especially in North American urban areas, securing the connectivity of transport to be independent of personal vehicles could be challenging. Further research for optimal scale and land use patterns of metropolitan areas under various urban conditions from the perspectives of economies, environment and quality of inhabitants is required.

![Image](image_url)

**Figure 7.** Share of population with access to public transport in urban core areas. Source: The OECD; EU, Access to Public Transport: Comparing a Selection of Medium-Sized and Large Cities; OECD Territorial Development Policy Committee: Paris, France, 2014.

While the findings of this study are encouraging, there are limitations to be addressed in the following research. First, the coverage of GTFS data across the globe has been increasing, yet the spatial coverage still concentrates on North America and Europe. As spatial scale and patterns of metropolitan areas among North America, Europe, and Asia are different, a wider and deeper comparison can be conducted with provision of GTFS data from Asian transit authorities later. Second, even though this study uses ambient population, there is still an issue of aggregate data and the assumption that the population in a cell is evenly distributed. This could be addressed by new data collection methods such as people-based tracking with ICT devices.

**Acknowledgments:** This study was conducted as a project, Access to Public Transport, in the OECD with European Commission (EC). This research was supported by Basic Science Research Programme through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2014R1A1A005367). This research was supported by a grant [MPSS-NH-2015-79] through the Natural Hazard Mitigation Research Group funded by Ministry of Public Safety and Security of Korean Government. This paper has previously been presented at the 8th International Conference of the International Forum on Urbanism (IFoU) on 22–24 June 2015, Incheon, Korea.

**Author Contributions:** Jinjoo Bok collected and analyzed data and wrote the original version of the manuscript. Youngsang Kwon finalized the manuscript and supplementary information. Both authors have read and approved the final manuscript.

**Conflicts of Interest:** The authors declare have no conflict of interest to declare.

**References**

1. Matulin, M.; Bosnjak, I.; Simunovic, L. Different Approaches to the Modal Split Calculation in Urban Areas; University of Zagreb: Zagreb, Croatia, 2009.
2. Lei, T.L.; Church, R.L. Mapping transit-based access: Integrating GIS, routes and schedules. *Int. J. Geogr. Inf. Sci.* 2010, 24, 283–304. [CrossRef]


7. Bhat, C.; Bricka, S.; Mondia, J.L.; Kapur, A.; Guo, J.; Sen, S. Metropolitan Area Transit Accessibility Analysis Tool; Centre for Transportation Research University of Texas: Austin, TX, USA, 2006; pp. 33–44.


33. Andresen, M.A. The ambient population and crime analysis. *Prof. Geogr.* **2011**, *63*, 193–212. [CrossRef]


© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).