

## Chapter III

# Combining Helsinki Region Travel Time Matrix with Lipas-database to analyse accessibility of sports facilities

Heittola, S., Koivisto, S., Ehnström, E. & Muukkonen, P.\*

[suvi.heittola@helsinki.fi](mailto:suvi.heittola@helsinki.fi), University of Helsinki  
[sonja.koivisto@helsinki.fi](mailto:sonja.koivisto@helsinki.fi), University of Helsinki  
[emil.ehnstrom@helsinki.fi](mailto:emil.ehnstrom@helsinki.fi), University of Helsinki  
[petteri.muukkonen@helsinki.fi](mailto:petteri.muukkonen@helsinki.fi), University of Helsinki

\* Corresponding author [petteri.muukkonen@helsinki.fi](mailto:petteri.muukkonen@helsinki.fi)

## Abstract

This project aims to simplify the process of connecting sports facility data from Lipas database (University of Jyväskylä, 2020a) with the Helsinki Travel Time Matrix (Accessibility research group, n.d.). The Lipas database contains spatial data over various sports facilities in Finland and the Helsinki Travel Time Matrix is a collection of files organised as a grid with information about travel times in the Helsinki Metropolitan Area. With a connection between these two data sets, it will be possible to see the travel times to a certain sports facility category in the Helsinki region. Establishing a link between these datasets can be useful for individuals, researchers and planners. With this toolpack one can easily select a sports facility category and get a TIFF raster returned, with the travel times included. The development of a toolpack has been made with Python programming language and it is available on GitHub (<https://github.com/petterimuukkonen/sportsfacilities>).

Keywords: accessibility; sport facility; travel time; Helsinki Metropolitan area; geodata merging; automatization

## Introduction

Accessibility to services and facilities is an important factor shaping the growth and spatial change in cities (Hasan et al. 2017). Spatial accessibility can be defined as the ability to reach goods, services and activities or as the degree to which a service or facility is accessible by as many people as possible (Litman 2010; Reggiani et al. 2015). The concept of accessibility has also physical, cultural, socio-economic and financial aspects but this study is focused on the geographical aspect (Karusisi et al. 2013).

Spatial accessibility to services can vary greatly within a metropolitan area and therefore the place of residence can either limit or enable our everyday activities (Higgs et al. 2015). In the case of sports facilities, the accessibility may affect the use rate of facilities and therefore have positive health impacts in the area (Karusisi et al. 2013). Furthermore, to prevent possible spatial marginalisation and health problem accumulation, it is important to ensure that all neighbourhoods have access to sports facilities and to investigate how the accessibility of facilities could be further improved with urban planning.

The accessibility research has developed greatly with the progress of Geographical Information Systems (GIS), tools and software (O'Sullivan et al. 2000). Many challenges still remain, since the accessibility varies according to the travel method, time of the day, area and rush hour patterns among other factors. Salonen and Toivonen (2013) have addressed this issue by creating comparable travel times for bike, public transport and car with a door-to-door approach that takes into account the rush hour, waiting and transfer times for public transport as well as time used for finding a parking place. The same approach is also used in this project to make the travel modes more comparable.

This project provides tools for accessibility research related to sports facilities in the Helsinki Metropolitan Area. Our aim is to automate a process of combining Finnish sport facility data (Lipas) with accessibility data of the Helsinki region area (Helsinki Region Travel Time Matrix). As a result of the process, a user gets a raster file of the accessibility of the chosen sport facilities in the Helsinki region and can visualise the results with static and interactive maps. This project is part of GIS project work course 2020 in University of Helsinki.

## **Data**

### ***Helsinki Travel Time Matrix and YKR grid***

Helsinki Region Travel Time Matrix 2018 (later HRTTM) is a set of text-files that consists of calculated travel times and distances from the Helsinki metropolitan area. The data is built on SYKE (Finnish Environmental Institute) YKR grid that has 13 231 individual grid cells, with width and height of 250 m. The travel times have been

calculated from each YKR grid cell centroid to each YKR grid cell centroid and saved into separate text-files. Each text-file represents travel times and distances from surrounding grid cells towards a single grid cell, so that the total count of text-files in this dataset is 13 231. For making the text files spatial data, the HRTTM data can be easily joined to a YKR grid file that has already been clipped to match the Helsinki Region area. The HRTTM dataset was calculated and collected in 2018 by MetropAccess-project and Accessibility Research Group in University of Helsinki and is the latest publication of Helsinki Region Travel Time Matrix datasets at the time. It is licensed under a Creative Commons Attribution 4.0 International license, and can be used freely (Accessibility research group, n.d.).

Travel times are calculated separately for four travelling modes; public transport, private car, cycling and walking (Accessibility research group, n.d.; Tenkanen et al. 2020). These four travel modes are furthermore divided into 10 different travel methods using different travel speeds and times of a day.

Travel times for public transport and private car are calculated using a door-to-door approach. This approach takes into consideration the entire journey from starting point to the destination in different travel modes, such as walking into a bus stop or parking lot and also possible waiting times on the way. Travel times are calculated in two different times of a day: morning rush hour (08:00-09:00) and midday (12:00-13:00). In addition to this, travel times for private car are also calculated based on existing speed limits and in public transport there is an option of choosing door-to-door approach with or without the possible waiting time at home before leaving (Tenkanen et al. 2020).

Since personal characteristics of a cyclist influence the travel speed vastly the travel times for cycling is divided into fast and slow cycling based on Strava network travel speed data averages. Additional minute is added into cycling times referring to unlocking and locking the bike. A static walking speed of 70 metres per minute is used in all travel modes (Tenkanen et al., 2020).

Each txt-file consists of 18 attributes: 1) from\_id, 2) to\_id, 3) walk\_t, 4) walk\_d, 5) bike\_f\_t, 6) bike\_s\_t, 7) bike\_d, 8) pt\_r\_tt, 9) pt\_r\_t, 10) pt\_r\_d, 11) pt\_m\_tt, 12) pt\_m\_t, 13) pt\_m\_d, 14) car\_r\_t, 15) car\_r\_d, 16) car\_m\_t, 17) car\_m\_d, 18) car\_sl\_t (Accessibility research group, n.d.). The first two attributes define from which YKRgrid cell to which YKR-grid cell the calculations have been taken from. The attributes

contain a YKR-ID for locating the exact grid cell. The last 16 attributes (nro 3 to 18) contain the calculated travel times and distances in different travel modes. In the attribute names walk stands for walking, bike for cycling, car for private car and pt for public transportation. The latter letter specifications are explained in table 1. Nodata values are presented as -1 in the dataset (Accessibility research group, n.d.).

**Table 1.** Specification of the letters in attribute field names in Helsinki Region Travel Time Matrix 2018 data.

Letters in attribute field names	Definition
t	Travel time in minutes
tt	Travel time in minutes for public transport with waiting time at home before leaving
d	Travel distance in meters
f	Fast speed for cycling (19 km/h)
s	Slow speed for cycling (12 km/h)
r	Rush hour (08:00–09:00, 29.01.2018)
m	Midday (12.00–13:00, 29.01.2018)
sl	Travel speed according to speed limits

### *Lipas sport facility data*

Lipas (<https://www.lipas.fi/etusivu>) is a national database of sport facilities in Finland. The database is managed by the Faculty of Sport and Health Sciences in the University of Jyväskylä and funded by Ministry of Education and Culture (University of Jyväskylä, 2020a). The data is uploaded to the database by municipality sport service workers or by private operators, such as outdoor unions and sport governing bodies (University of Jyväskylä, 2020b). Lipas data can be used through Liikuntapaikat.fi -application, by downloading ready-made datasets or by using an open source web map service (WMS), a web feature service (WFS) or a REST-interface (University of Jyväskylä, 2020c).

Lipas data is open source data that has been licenced by Creative Commons Attribution 4.0 International (CC BY 4.0) (University of Jyväskylä, 2020b).

The Lipas database contains information on sport facilities that includes maintained sport places, outdoor routes and parks (University of Jyväskylä, 2020a). A sport facility has to be publicly accessed, maintained regularly and equipped appropriately to be added into the database. A sport facility can also be for example a guiding point, maintenance building or an outdoor fireplace that is related to sports or outdoor activities. All in all, there were over 37 000 sport facilities in the dataset in spring 2019 (University of Jyväskylä, 2020b).

The sport facilities are grouped into sport facility types. There are eight (8) main types of sport facilities in the database (University of Jyväskylä, 2020b):

- (1) Outdoor places and services (Virkistyskohteet ja palvelut in Finnish)
- (2) Outdoor courts and sport parks (Ulkokentät ja liikuntapuistot)
- (3) Indoor sport places (Sisäliikuntapaikat)
- (4) Water sport places (Vesiliikuntapaikat)
- (5) Terrain/outdoor sport places (Maastoliikuntapaikat)
- (6) Boating, aviation and motor sports (Veneily, ilmailu ja moottoriturheilu)
- (7) Animal sports (Eläinurheilu)
- (8) Maintenance buildings (Huoltorakennukset)

These main types have been further on divided into more specific sport facility types. Many of the types have specific information that is only relevant in that sport facility. Therefore, these sport types differ from each other also based on attribute fields. However, there are few attribute fields that exist in all of these type groups. These are for example the sport facility type name in Finnish, Swedish and English, sport facility type code, the actual name of the facility and so on.

## Methods and workflow for combining datasets

The Lipas sport facility data will be fetched from its open source Web Feature Service (WFS). This way it is possible to use Lipas data without storing it locally and always use the most updated data. The Lipas database includes features from all geometry types; points, lines and polygons. However, area and line features are somewhat problematic when it comes to accessibility analyses. One cannot be exactly sure from which point a person can truly access for example to a hiking route or a forest park area. Therefore, in this project we will be using only point based sport facility data.

The HRTTM 2018 and YKR grid data will be downloaded to a local repository and used from there through the process, since the data are stable releases and aren't updated regularly. The HRTTM and YKR grid data are openly available and can be downloaded from the data providers' website.

The actual process of combining Lipas data to HRTTM data consists of six separate steps (figure 1). First the Lipas sport facility data we will be obtained with a WFS request (step 1). After this Lipas data will be spatially joined to a YKR grid file to understand in which grid cells the sport facilities lie into (step 2). By this we will know the correct YKR grid cell ids that will help us to find out which HRTTM text-files are relevant in the outcoming maps.

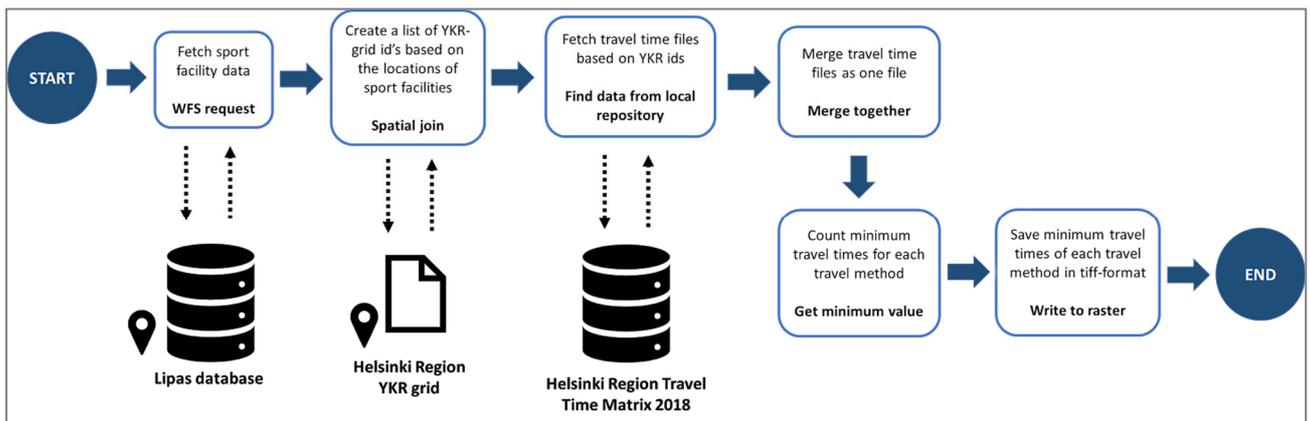


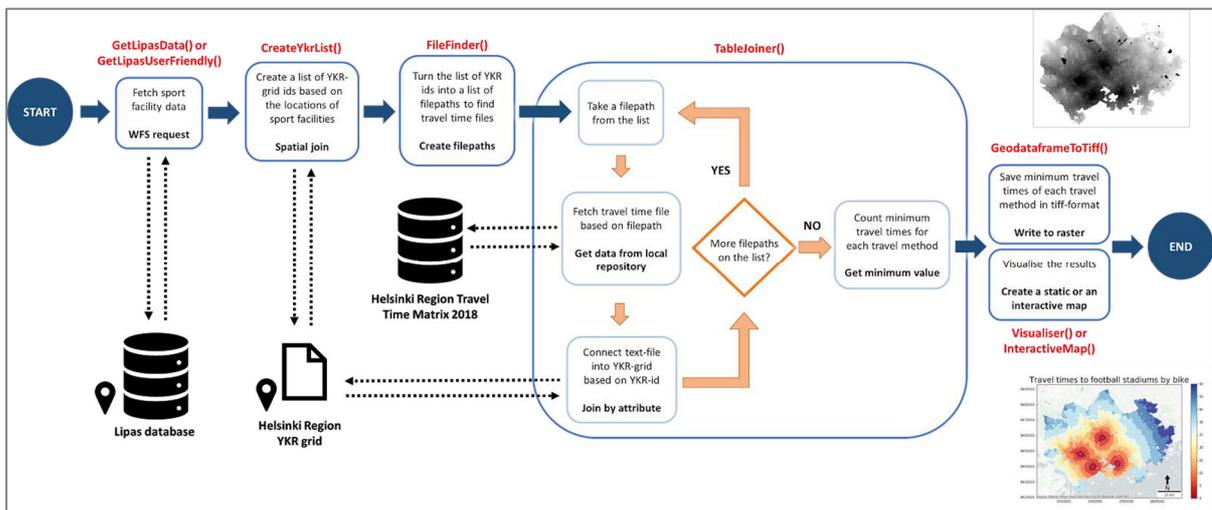
Figure 2. Simple workflow of the planned combining process.

When we have a list of YKR grid ids that correspond to the locations of the sport facilities we can obtain the correct HRTTM text-files from a local repository (step 3). If the chosen sport facility type has more than one sport facility there will most likely be more than one HRTTM file. These HRTTM files need to be merged as one file (step 4) and then the minimum travel times to any of the sport facilities can be calculated (step 5). After the minimum travel times have been calculated the data can be written into a raster file, visualized and used for example in research purposes (step 6).

We'll use Python programming language to build the automated process of combining the data and GitHub-platform as a version control service and shared platform for developing the code. GitHub will also be used in storing, developing and sharing the automated process further on.

## Results

The final product of this project is a tool pack consisting of Python code functions that can be used for combining HRTTM data and Lipas data. The tool pack also includes functions that can be used for visualizing the results easily on a map. The tool pack includes eight functions (figure 2) in total that aim to simplify the process of combining HRTTM data with Lipas data. As the final output the tool pack can return the combination of the data as a TIFF raster file that includes travel times to chosen sports facilities with a particular travelling method. This raster file can then be used for further analysis or research and also be visualised in various GIS softwares.



**Figure 2.** Workflow and the functions of the created tool pack.

The tool pack can be used by anyone, but since the tool pack is based on Python programming language, the assumption is, that it is most useful for developers, researchers and other people that have at least a brief understanding of programming. This tool pack is made with the geographical extent of the Helsinki Metropolitan Area and is therefore only useful for people interested in the particular area.

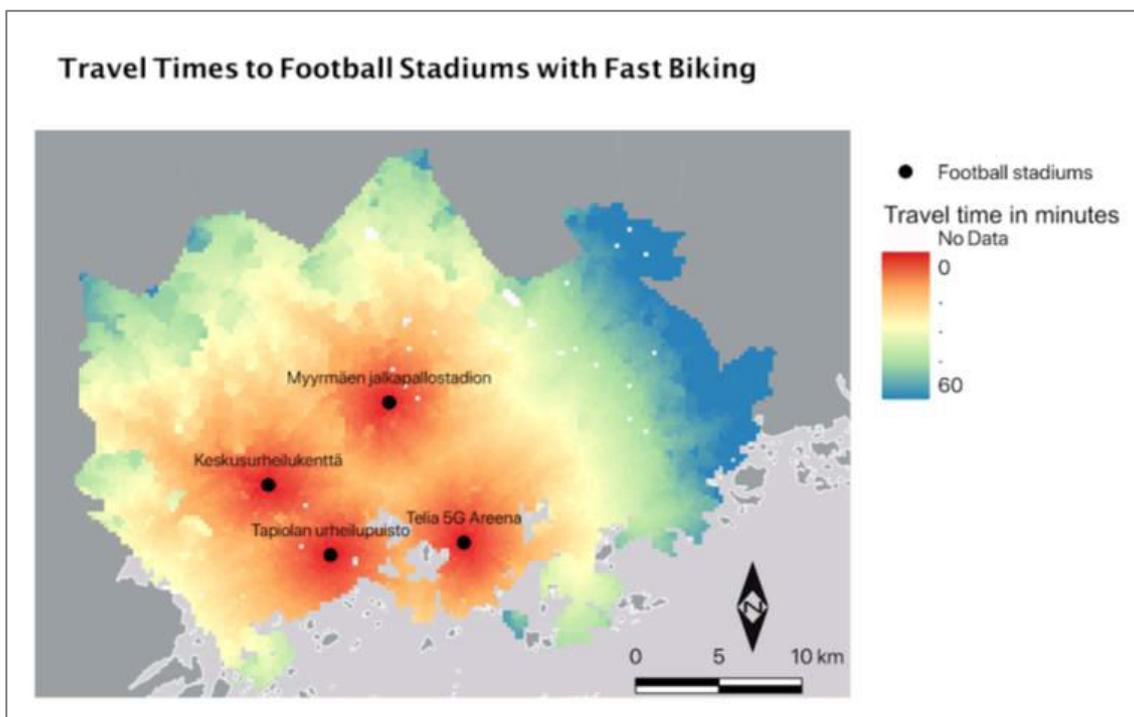
After downloading the HRTTM and YKR grid data from the data provider's website you can start using the functions found on Github. The sport facility data will be fetched from the Lipas database with a WFS request. This can be done with two functions: `GetLipasData()` or `GetLipasUserFriendly()`. By calling `GetLipasData()` – function with two parameters, the sport facility type code and type name, the user will do a WFS request to the Lipas server, which will fetch the data. The type codes and type names can be found in a designated list in the GitHub repository or from Lipas database webpage (University of Jyväskylä, 2020d). As an example, one could choose to call `GetLipasData("1350", "Jalkapallostadion")` for fetching football stadiums. The `GetLipasUserFriendly()` -function is made for a more user-friendly approach. To use this function the user doesn't need to know in advance which type code or exact type name the sport facility has in Lipas database. As the function is called, it will list all available sport facility types for the user and the user can choose the one (s)he wants.

The fetched Lipas data needs to be spatially joined with the YKR grid, before it can be merged to the travel time matrix data. This can be done by using a function called `CreateYkrList()`. The function creates a list of YKR grid cell ids that have a sports facility inside them. As an input parameter user inserts the Lipas data that was fetched from the Lipas server. After the YKR id list is created a function called `FileFinder()` will turn the list of ids into a list of filepaths to the correct travel time files in the HRTTM dataset.

`TableJoiner()` function will fetch all the HRTTM text-files found based on the filepath list and turns the text-files into a geodataframe by joining the files into the YKR grid. After the files are joined as one, minimum travel times from each cell to the sport facility are calculated and a geodataframe is returned to the user.

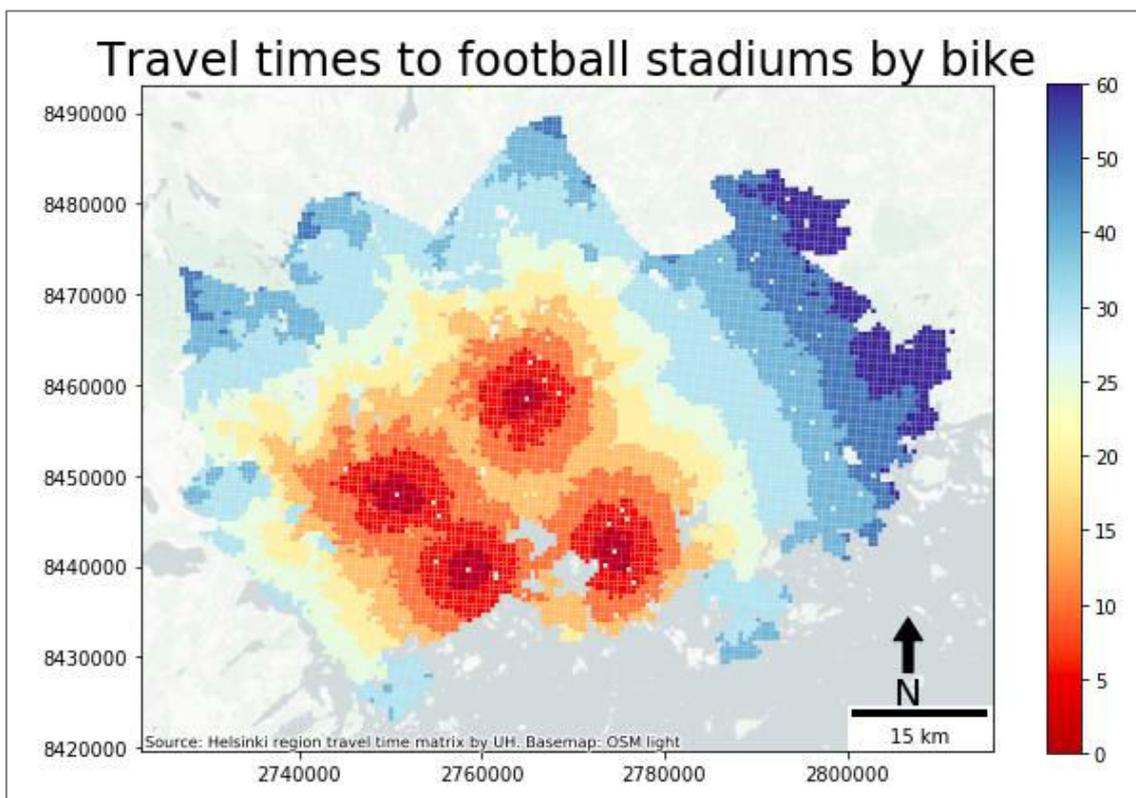
After the table joining process, the data needs to be written into a raster file, so that it's easier to use. This is done by the function `GeodataframeToTIFF()` that takes a geodataframe (recently created with the `TableJoiner()`), sport facility type name and type code as input parameters. The function turns the input geodataframe of minimum

travel times into raster format and returns a set of raster TIFF files for each of the ten different travel modes in HRTTM data. These raster files can then be easily used to visualize the accessibility of the sport facilities (figure 3).



**Figure 4.** Simple visualisation of the raster output file in QGIS.

The Visualiser() function requires three parameters: a geodataframe with minimum travel times, the column name of the chosen travel method and the sport facility type name. The sport facility type name is only used for the output map title. Since this function has been designed to work only in English, it is recommended to insert the sport facility type name in English. As an output Visualiser () generates a map presentation as a png file and places it in the output folder. The map visualises the minimum travel times to the closest sport facility in chosen travel mode. The output file is a static image and is therefore easy to use in text documents and on paper (figure 4).



**Figure 4.** An example of what kind of output the Visualiser() function creates for travel times to football stadiums by bike.

A more suitable map for the web can be made with the function called InteractiveMap(). The function requires two parameters: one geodataframe and a column name (in other words the travel mode you wish to display). The InteractiveMap() returns an html file that contains an interactive choropleth map with the correct travel times for the chosen sports facility type. All the functions can be found on the GitHub page <https://github.com/petterimuukkonen/sportsfacilities>.

## Discussion

The tool pack we present serves functional purposes, however, for someone who is not familiar with programming, learning how to use the tools can be overwhelming. Even for people who are familiar with Python programming probably need to refer to the documentation and demo material we have made in order to effectively use the functions. We see a need for a more user-friendly approach to our problem, where anyone could choose any set of sports facilities to be reached. This would make our work more

accessible to the masses and our groundwork would become more useful. A solution to this would be producing a geodatabase containing the accessibility rasters for all travel methods for all sports facility types. This would be a kind of online library where all the files are stored and can be freely used for research purposes.

Currently, the geographical extent for using the tools we have created is quite limited. We are using the HRTTM as an input, and this type of accessibility data has not been created for different areas with the same format. Furthermore, the production of this type of accessibility data requires time resources and funding. Therefore, the approach is not feasible for extensive areas, at least with the current level of technology. If travel time matrices were created for other city regions in Finland, the tools could be easily extended to work for those areas as well. This project would not have been possible without up-to-date and quality data sources like HRTTM and Lipas database. If other countries provide and maintain similar databases, comparative research could be done and this can improve the understanding of spatial accessibility of sport facilities globally and the consequences.

We acknowledge that there are many possible points of improvement in our work. For a 2.0 version, if one would be made, we would add many functionalities which include at least the following:

In this version, the polygon and line features, like outdoor areas or orienteering terrains and jogging paths or ski tracks, are just filtered out. In the next version, the travel times could be calculated to the closest point of the polygon or polyline if it can be entered from any point. Another option would be to locate the enter points of routes or park areas and such, if that is even possible.

Currently, the user can only choose one sport facility subtype (like football fields) when fetching the data without modifying the function. For 2.0 version, the Lipas data fetching functions could be developed to fetch all subtype features inside a chosen sport facility maintype (like ballsport fields) or to choose multiple separate sport facility subtypes to be fetched at once, if the user wants so.

Currently, there is a slight bug in the HRTTM data, which results in NoData value in the YKR grid cell where the travel times have been calculated to. Therefore in some outcome raster files there are NoData values (-1) in the raster cells which represent the locations of sport facilities. These NoData values could be manually turned into 0 values, in order to distinguish it to be a destination point (where travel time is 0) and not actually missing data.

The Visualiser() function could be upgraded to have multiple languages and choose the language according to user input. A challenge here is the conjugation of words in Finnish.

Ideally, we would also host the travel time data (HRTTM and YKR grid) on an online platform where it is ready to be used with our demo notebooks and functions so that every user would not have to download and extract the files which can take some time.

## Conclusions

Accessibility of sports facilities is an important topic that can affect the activity levels and even health of urban dwellers (Higgs et al. 2015; Karusisi et al. 2013). With this project, we have provided a toolset for further accessibility research by writing functions to combine the accessibility data and the sports facility data. With these tools, the user is able to choose facilities of their interest and assess the accessibility of those by creating maps and raster files for further processing and analysis in GIS softwares. Furthermore, different time considerations like rush hour and waiting time at home for public transport have been taken into account for the different travel methods and this makes them more realistic and comparable with each other (Accessibility research group, n.d.).

The functions we created can be run for different sport facilities inputs and the results can be gathered to an accessibility geodatabase that could be available online for other users. This was one of our goals, but we could not see it through due to restrictions during the pandemic. Freely accessible database could facilitate more accessibility research and make the results available for all, not just for those who handle the basics of Python.

## References

- Accessibility research group (n.d.). Helsinki Region Travel Time Matrix (2018). University of Helsinki. <https://blogs.helsinki.fi/accessibility/helsinki-region-travel-timematrix-2018/> Read 4.3.2020.
- Hasan, S., Wang, X., Khoo, Y.B. & Foliente, G. (2017). Accessibility and socio-economic development of human settlements. *PLoS ONE* 12(6). <https://doi.org/10.1371/journal.pone.0179620>

- Higgs, G., Langford, M. & Norman, P. (2015). Accessibility to sport facilities in Wales: A GIS-based analysis of socio-economic variations in provision. *Geoforum* 62: 105–120. <https://doi.org/10.1016/j.geoforum.2015.04.010>
- Karusisi, N., Thomas, F., Méline, J. & Chaix B. (2013). Spatial accessibility to specific sport facilities and corresponding sport practice: the RECORD Study. *International Journal of Behavioral Nutrition and Physical Activity* 10, 48. <https://doi.org/10.1186/1479-5868-10-48>
- Litman, T. (2010). Accessibility. A Dictionary of Transport Analysis. 2010. pp. 1–3.
- O'Sullivan, D., Morrison, A. & Shearer, J. (2000). Using desktop GIS for the investigation of accessibility by public transport: an isochrone approach. *International Journal of Geographical Information Science* 14(1), 85–104. <https://doi.org/10.1080/136588100240976>
- Reggiani, A., Nijkamp, P. & Lanzi, D. (2015) Transport resilience and vulnerability: The role of connectivity. *Transportation Research Part A: Policy and Practice* 81, 4–15. <https://doi.org/10.1016/j.tra.2014.12.012>
- Salonen, M. & Toivonen, T. (2013). Modelling travel time in urban networks: comparable measures for private car and public transport. *Journal of Transport Geography* 31, 143–153. <https://doi.org/10.1016/j.jtrangeo.2013.06.011>
- Tenkanen, H. & Toivonen, T. (2020). Longitudinal spatial dataset on travel times and distances by different travel modes in Helsinki Region. *Scientific Data* 7, 77. <https://doi.org/10.1038/s41597-020-0413-y>
- University of Jyväskylä (2020a). Lipas Liikuntapaikat.fi. <https://www.jyu.fi/sport/fi/yhteistyo/lipas-liikuntapaikat.fi>. Read 17.3.2020.
- University of Jyväskylä (2020b). Lipas-järjestelmän esittely. <https://www.jyu.fi/sport/fi/yhteistyo/lipas-liikuntapaikat.fi/esittely-2>. Read 17.3.2020.
- University of Jyväskylä (2020c). Rajapinnat ja ladattavat aineistot. <https://www.jyu.fi/sport/fi/yhteistyo/lipas-liikuntapaikat.fi/rajapinnat-jaladattavat-aineistot>. Read 17.3.2020.
- University of Jyväskylä (2020d). Lipas 2019 (Lipas 2.0) käyttöohjeet. <https://www.jyu.fi/sport/fi/yhteistyo/lipas-liikuntapaikat.fi/lipas-2019-2-0>. Read 29.3.2020.