Comparison of management strategies for the charging schedule and all-electric operation of a plug-in hybrid-electric bi-articulated bus fleet

Supplementary material:

Data preparation process

# Authors

Dennis Dreier 1,\*,\*\*, Björn Rudin 2,\*\*, Mark Howells 3,4,\*\*

1 KTH Royal Institute of Technology, Department of Energy Technology, Stockholm, Sweden

2 Combitech AB, Linköping, Sweden

3 Geography and Environment, School of Social Sciences, Loughborough University, Loughborough, UK

4 Center for Environmental Policy, Imperial College London, London, UK

\* Corresponding author. E-Mail address: ddreier@kth.se, Postal address: KTH Industrial Engineering and Management, Department of Energy Technology, Brinellvägen 68, 100 44 Stockholm, Sweden

\*\* ORCID numbers: Dennis Dreier (0000-0002-0437-2093), Björn Rudin (0000-0002-2619-366X), Mark Howells (0000-0001-6419-4957)

# Data preparation

This section is addressed to whom is interested in using the bus operations data for the public bus transport system in the city of Curitiba in Southern Brazil available on the open data platform at Ref. (UFPR 2017). The following information explains the the content of the bus operation files and how to prepare them so that they can be used as input data for the RTO-model.

1. Bus operation data sets on Curitiba’s open data platform are named as ”yyyy\_mm\_dd\_veiculos.json.xz” (year\_month\_day). The file name extension “.json” shows that data is stored in the json format. A bus operation data file contains the following information:
	1. ”COD”: bus route number
	2. ”LAT”: latitude
	3. ”LON”: longitude
	4. ”VEIC”: bus identification number
	5. ”HORA”: timestamp of measurement (date and time)
2. Errors in the data sets were excluded such as incomplete data points (i.e. if one or more of the above information were missing). Then, the following information were generated:
3. Interpolation of the data points according to the decided duration of one time step ****** in the simulations.
4. Calculation of the time difference between data points i and i+1: 
	1. 
5. Calculation of the distance between data points i and i+1: 
	1. The Haversine formula was used to calculate the distance between two consecutive GPS coordinates of a bus (great circle radius of the earth: 6371 km). Various code implementations of this function are available at Ref. (Rosetta Code 2018)
	2. 
6. Calculation of the average speed between data points i and i+1: 
	1. step function for the average speed between data points
	2. 
7. Calculation of the average acceleration between data points i and i+1: 
	1. step function for the average acceleration between data points
	2. 
8. Determination of the elevation at data point i
	1. The elevation at the GPS coordinates was determined by using SRTM3 (i.e. Shuttle Radar Topography Mission) data and downloading the necessary hgt files
9. Calculation of the elevation difference between data points i and i+1: 
	1. 
10. Calculation of the grade between data points i and i+1: s
	1. 
11. Determination whether a bus was at a terminal station or not for data point i:
	1. Use data file named as ”yyyy\_mm\_dd\_pontosLinha.json.xz” (year\_month\_day) to obtain the GPS position of the bus stations in Curitiba. This file contains information about the bus stations such as:
		1. ”COD”: bus route number
		2. ”LAT”: latitude
		3. ”LON”: longitude
		4. ”NOME”: name of the bus station
		5. ”SENTIDO”: name of end station
		6. ”SEQ”: ascending sequence number of bus stations on a bus route
	2. Calculate local minima between the distances of the GPS coordinates of the bus stations in the file ”yyyy\_mm\_dd\_pontosLinha.json.xz” and GPS positions of the buses in the file ”yyyy\_mm\_dd\_veiculos.json.xz”. Set a cut-off value, i.e. distances below or equal to the cut-off value indicate that a bus was at the bus station. Note: A terminal station (as used in this study) has a relatively large spatial area. Therefore, even if a bus was at a terminal station, the calculated distance between the centre of terminal station (GPS point of terminal station) and the GPS position of a bus at a data point can be relatively large. Thus, on one hand, setting a small/short cut-off limit can lead to the result that a bus is not considered as being at a bus station, although it would actually be there. On the other hand, too large/long cut-off values can lead to result that a bus is considered as already being at a bus station, although it would actually not be there. In this study, a cut-off value of 90 m was chosen based on a previous field study trip by the first author of this study. (Note: some field study information are available in references (Dreier et al. 2018; Dreier 2015)).
	3. The same approach led to the determination of the actual arrival time and departure time of a bus from a terminal station.

# References

 Dreier, D.: Assessing the potential of fuel saving and emissions reduction of the bus rapid transit system in Curitiba, Brazil, (2015)

 Dreier, D., Silveira, S., Khatiwada, D., Fonseca, K.V.O., Nieweglowski, R., Schepanski, R.: Well-to-Wheel analysis of fossil energy use and greenhouse gas emissions for conventional, hybrid-electric and plug-in hybrid-electric city buses in the BRT system in Curitiba, Brazil. Transportation Research Part D: Transport and Environment. 58, 122–138 (2018). doi:10.1016/j.trd.2017.10.015

 Rosetta Code: Various code implementations of the Haversine formula, https://rosettacode.org/wiki/Haversine\_formula

 UFPR: Open data online platform of the UFPR – Federal University of Paraná, bus transport system data sets, http://dadosabertos.c3sl.ufpr.br/curitibaurbs/