ZTBus: A Dataset of 1000+ Complete, Second-Resolved Driving Missions of Inner-City Transit Buses

----------------------------------------------------------------------------------------------------------

Abstract:

This repository contains the Zurich Transit Bus (ZTBus) dataset, which consists of data recorded during

driving missions of electric city buses in Zurich, Switzerland. The data was collected over several years

on two trolley buses as part of multiple research projects. It involves more than a thousand missions

spanning across all seasons, each mission usually covering a full day of real operation. The ZTBus

dataset contains detailed information on the vehicle’s power demand, propulsion system, odometry, global

position, ambient temperature, door openings, number of passengers, dispatch patterns within the public

transportation network, etc. All signals are synchronized in time and include an absolute timestamp in

tabular form. The dataset can be used as a foundation for a variety of studies and analyses. For example,

the data can serve as a basis for simulations to estimate the performance of different public transit

vehicle types, or to evaluate and optimize control strategies of hybrid electric vehicles. Furthermore,

numerous influencing factors on vehicle operation, such as traffic, passenger volume, etc., can be

analyzed in detail.

This dataset is supplemented by the accompanying data descriptor paper:

> F. Widmer, A. Ritter, and C. H. Onder, "ZTBus: A Large Dataset of Time-Resolved City Bus Driving

Missions," Scientific Data, vol. 10, no. 1. Springer Science and Business Media LLC, Oct. 10, 2023.

https://doi.org/10.1038/s41597-023-02600-6.

Please reference this publication when using this data.

----------------------------------------------------------------------------------------------------------

Authors responsible for the collection and curation of the data:

Fabio Widmer (\*)

fawidmer@idsc.mavt.ethz.ch

Institute for Dynamic Systems and Control, ETH Zurich, 8092 Zurich, Switzerland

https://orcid.org/0000-0003-0604-3856

Dr. Andreas Ritter (\*)

anritter@idsc.mavt.ethz.ch

Institute for Dynamic Systems and Control, ETH Zurich, 8092 Zurich, Switzerland

https://orcid.org/0000-0002-7073-7118

Prof. Dr. Christopher H. Onder

onder@idsc.mavt.ethz.ch

Institute for Dynamic Systems and Control, ETH Zurich, 8092 Zurich, Switzerland

(\*): These authors contributed equally.

----------------------------------------------------------------------------------------------------------

General information about the data:

Period of collection: April 2019 - December 2022

Location of collection: Zurich, Switzerland

License: CC-BY-4.0

How to cite: Please reference the accompanying data descriptor paper:

> F. Widmer, A. Ritter, and C. H. Onder, "ZTBus: A Large Dataset of Time-Resolved City Bus Driving

Missions," Scientific Data, vol. 10, no. 1. Springer Science and Business Media LLC, Oct. 10, 2023.

https://doi.org/10.1038/s41597-023-02600-6.

The data was recorded on two trolley buses during regular operation by Verkehrsbetriebe Zürich (VBZ).

Both are "HESS lighTram® 19 DC" buses, which are single-articulated, have an overall length of about 19

m, a curb weight of about 19 t, and a maximum passenger capacity of about 160. They are equipped with

traction batteries, which allow them to run for a few kilometers without the overhead power grid. The

dataset covers the operation of the buses on various bus routes in Zurich's public transportation network.

More information about the data collection and curation pipeline can be found in the accompanying data

descriptor paper mentioned above.

----------------------------------------------------------------------------------------------------------

Research project funding:

> Carrosserie HESS AG

> Verkehrsbetriebe Zürich (VBZ)

> Swiss Federal Office of Energy (SFOE): Contract numbers SI/501321, SI/501979, SI/502235

----------------------------------------------------------------------------------------------------------

Publications based on this data:

> A. Ritter, F. Widmer, J. W. Niam, P. Elbert, and C. H. Onder, "Real-Time Graph Construction Algorithm

for Probabilistic Predictions in Vehicular Applications," in IEEE Transactions on Vehicular Technology,

vol. 70, no. 6, June 2021, https://doi.org/10.1109/TVT.2021.3077063.

> A. Ritter, F. Widmer, B. Vetterli, and C. H. Onder, "Optimization-based online estimation of vehicle

mass and road grade: Theoretical analysis and experimental validation", in Mechatronics, Volume 80, 2021,

https://doi.org/10.1016/j.mechatronics.2021.102663.

> A. Ritter, F. Widmer, P. Duhr, and C. H. Onder, "Long-term stochastic model predictive control for the

energy management of hybrid electric vehicles using Pontryagin's minimum principle and scenario-based

optimization", in Applied Energy, Volume 322, 2022, https://doi.org/10.1016/j.apenergy.2022.119192.

> F. Widmer, A. Ritter, P. Duhr, and C. H. Onder, "Battery lifetime extension through optimal design and

control of traction and heating systems in hybrid drivetrains", in eTransportation, Volume 14, 2022,

https://doi.org/10.1016/j.etran.2022.100196.

> F. Widmer, A. Ritter, J. Ritzmann, D. Gerber, and C. H. Onder, "Battery Health Target Tracking for

HEVs: Closed-Loop Control Approach, Simulation Framework, and Reference Trajectory Optimization", in

eTransportation, Volume 17, 2023, https://doi.org/10.1016/j.etran.2023.100244.

> F. Widmer, A. Ritter, M. Achermann, F. Büeler, J. Bagajo, and C. H. Onder, "Highly Efficient Year-Round

Energy and Comfort Optimization of HVAC Systems in Electric City Buses", preprint presented at the IFAC

World Congress, 2023, https://doi.org/10.48550/arXiv.2303.00571.

----------------------------------------------------------------------------------------------------------

File description:

The ZTBus dataset is organized in two different types of comma-separated values (CSV) text files, which

are described below. Furthermore, for convenience, a sample Matlab script is provided that allows to load

and visualize some parts of the data.

Data files:

The names of the files that describe the individual driving missions are based on the vehicle

identification number (either 183 or 208) and the time period in which the data was collected. For

example, the data collected on the bus numbered 183 between 16 Oct 2019 02:52:43 and 16 Oct 2019

07:10:12, both given in UTC, is available in the following file:

B183\_2019-10-16\_02-52-43\_2019-10-16\_07-10-12.csv

The ZTBus dataset consists of 1409 driving missions, each of which is described in a separate CSV file.

All files have the same structure and format, where the first row contains the headers of the

corresponding columns and the remaining rows describe the set of data samples recorded at a specific

moment in time. This time index is represented in the first column as absolute UTC time, expressed

according to ISO 8601.

For convenience, the files are provided in a bundled format in the form of ZIP files:

> ZTBus.zip: Contains the entire ZTBus dataset, without compression to ensure long-time accessibility.

> ZTBus\_compressed.zip: Contains the entire ZTBus dataset in a compressed file.

> ZTBus\_samples.zip: Contains a random selection of 10 data files.

metaData.csv

This file contains metadata of the driving missions in a tabular form. The first row contains the headers

of the corresponding columns. The remaining rows contain metadata of the driving missions, indexed via

the corresponding file name in the first column.

loadAndVisualizeData.m

This sample Matlab code is added to the code repository for convenience. It is subjecto to the GNU

General Public License version 3 (GPLv3). It can be used to load parts of the data and recreate the

figures shown in the accompanying data descriptor paper. This code can serve as a starting point for own

analyses. It has been developed with Matlab version 9.12 (R2022a) and does not require any specialized

toolboxes.

----------------------------------------------------------------------------------------------------------

Detailed description of the data:

The data was recorded at different frequencies by various on-board devices. For convenience, we provide

the data in a synchronized form. Unless specified otherwise, "NaN" is used to represent unavailable data.

The columns of the data files are explained in the following:

> time\_iso (datetime): The absolute UTC time, expressed according to ISO 8601.

> time\_unix (integer) [s]: The Unix timestamp, i.e., the number of seconds passed since 00:00:00 UTC on

1970-01-01.

> electric\_powerDemand (float) [W]: The overall electric power demand of the vehicle. The values include

the power demand of the traction motors (which can be negative during recuperation phases) and all

auxiliary power consumers like HVAC, air compressor, lighting, infotainment systems, etc.

> gnss\_altitude (float) [m]: The altitude above sea level measured by the GNSS sensor.

> gnss\_course (float) [rad]: The course (heading) provided by the GNSS sensor. Traveling north is

represented by a course of 0 rad or 2\*pi rad. Traveling east is represented by a value of pi/2 rad. When

the bus is not moving, this estimate is held constant by the GNSS sensor on bus 183 and set to zero on

bus 208.

> gnss\_latitude (float) [rad]: The latitude on WGS 84 measured by the GNSS sensor.

> gnss\_longitude (float) [rad]: The longitude on WGS 84 measured by the GNSS sensor.

> itcs\_busRoute (string): The VBZ bus route name, provided by the ITCS. For unavailable data, we use a

dash ("-").

> itcs\_numberOfPassengers (float) [-]: The estimated number of passengers on the bus, measured with

infrared-based passenger counting systems.

> itcs\_stopName (string): The name of the bus stop served, provided by the ITCS. For unavailable data, we

use a dash ("-").

> odometry\_articulationAngle (float) [rad]: The angle of the pivoting joint (articulation). A positive

angle is observed in a right turn.

> odometry\_steeringAngle (float) [rad]: The angle of the front wheels relative to the vehicle body. A

positive angle is observed in a left turn.

> odometry\_vehicleSpeed (float) [m/s]: The vehicle speed. This value is based on the rotational velocity

of the drive shaft of the middle axis and a compound transmission ratio.

> odometry\_wheelSpeed\_fl (float) [m/s]: The wheel speed of the front left wheel, measured using a wheel

encoder. Due to the nature of the measurement, it is not reliable below approximately 1.5 m/s.

> odometry\_wheelSpeed\_fr (float) [m/s]: The wheel speed of the front right wheel, measured using a wheel

encoder. Due to the nature of the measurement, it is not reliable below approximately 1.5 m/s.

> odometry\_wheelSpeed\_ml (float) [m/s]: The wheel speed of the middle left wheel, measured using a wheel

encoder. Due to the nature of the measurement, it is not reliable below approximately 1.5 m/s.

> odometry\_wheelSpeed\_mr (float) [m/s]: The wheel speed of the middle right wheel, measured using a wheel

encoder. Due to the nature of the measurement, it is not reliable below approximately 1.5 m/s.

> odometry\_wheelSpeed\_rl (float) [m/s]: The wheel speed of the rear left wheel, measured using a wheel

encoder. Due to the nature of the measurement, it is not reliable below approximately 1.5 m/s.

> odometry\_wheelSpeed\_rr (float) [m/s]: The wheel speed of the rear right wheel, measured using a wheel

encoder. Due to the nature of the measurement, it is not reliable below approximately 1.5 m/s.

> status\_doorIsOpen (boolean): A binary flag indicating whether at least one door is open.

> status\_gridIsAvailable (boolean): A binary flag indicating whether the current collector of the trolley

bus is connected to the overhead grid.

> status\_haltBrakeIsActive (boolean): A binary flag indicating whether the halt brake is active. This

brake is automatically activated whenever the bus stands still.

> status\_parkBrakeIsActive (boolean): A binary flag indicating whether the park brake is active. This

fail-safe brake works also if the vehicle is not powered. It can be manually activated by the driver,

e.g., for longer standstill periods.

> temperature\_ambient (float) [K]: The measured ambient temperature. The temperature sensor is located

behind the front bumper. Due to the sensor resolution, it is only available in increments of 1 K.

> traction\_brakePressure (float) [Pa]: The mean pressure in the friction braking lines.

> traction\_tractionForce (float) [N]: An estimate of the overall traction force provided by the two

traction motors. These values are based on an estimate of the traction torque provided by the two

traction motors and a compound transmission ratio.

----------------------------------------------------------------------------------------------------------

Detailed description of the metadata:

The metadata consists of one row for each of the driving missions. The corresponding columns are

explained below:

> name (string): The name of the data file represented by the corresponding row of the metadata table.

> busNumber (integer) [-]: The number of the bus on which the data was recorded (183 or 208).

> startTime\_iso (datetime): The absolute UTC time of the start of the driving mission, expressed

according to ISO 8601.

> startTime\_unix (integer) [s]: The Unix timestamp of the start of the driving mission, in seconds since

00:00:00 UTC on 1970-01-01.

> endTime\_iso (datetime): The absolute UTC time of the end of the driving mission, expressed according to

ISO 8601.

> endTime\_unix (integer) [s]: The Unix timestamp of the end of the driving mission, in seconds since

00:00:00 UTC on 1970-01-01.

> drivenDistance (float) [m]: The distance covered during the entire driving mission. This value is

calculated via trapezoidal integration of the time series odometry\_vehicleSpeed.

> busRoute (string): The most frequently occurring value of the time series itcs\_busRoute.

> energyConsumption (float) [J]: The overall electric energy consumption of the vehicle during the entire

driving mission. This value is calculated via trapezoidal integration of the time series

electric\_powerDemand.

> itcs\_numberOfPassengers\_mean (float) [-]: The average number of passengers in the bus, i.e., the mean

value of the time series itcs\_numberOfPassengers, where NaN values are omitted.

> itcs\_numberOfPassengers\_min (float) [-]: The minimum number of passengers in the bus, i.e., the minimum

value of the time series itcs\_numberOfPassengers, where NaN values are omitted.

> itcs\_numberOfPassengers\_max (float) [-]: The maximum number of passengers in the bus, i.e., the maximum

value of the time series itcs\_numberOfPassengers, where NaN values are omitted.

> status\_gridIsAvailable\_mean (float) [-]: The fraction of time the bus is connected to the overhead

grid, i.e., the mean value of the time series status\_gridIsAvailable.

> temperature\_ambient\_mean (float) [K]: The average ambient temperature, i.e., the mean value of the time

series temperature\_ambient, where NaN values are omitted.

> temperature\_ambient\_min (float) [K]: The minimum ambient temperature, i.e., the minimum value of the

time series temperature\_ambient, where NaN values are omitted.

> temperature\_ambient\_max (float) [K]: The maximum ambient temperature, i.e., the maximum value of the

time series temperature\_ambient, where NaN values are omitted.