



TUMDOT–MUC: Data Collection and Processing of Multimodal Trajectories Collected by Aerial Drones

Alexander Kutsch¹ · Martin Margreiter¹ · Klaus Bogenberger¹

Received: 5 February 2024 / Revised: 26 June 2024 / Accepted: 28 June 2024 / Published online: 6 July 2024
© The Author(s) 2024

Abstract

Currently available trajectory data sets undoubtedly provide valuable insights into traffic events, the behavior of road users and traffic flow theory, thus enabling a wide range of applications. However, there are still shortcomings that need to be addressed: (i) the continuous temporal recording (ii) of a coherent area covering several intersections (iii) with the detection of all road users, including pedestrians and cyclists. Therefore, this study focuses on the design of a large-scale aerial drone observation in the city of Munich, Germany, as well as the processing steps and the description of the resulting data set. Using twelve camera-equipped, unmanned aerial drones, the observation monitored an inner urban road section with a length of 700 m continuously for several hours during the afternoon peak hours on two working days. The trajectories of all road users were then extracted from the videos and post-processed in order to obtain a coherent and accurate data set. The resulting trajectories contain information on the category, dimensions, location, velocity, acceleration and orientation of each road user at each frame, merged continuously in time and space across several drone observation areas and subsequent time slots. The data, therefore, includes various interactions between different modes of motorized traffic and active mobility users like pedestrians and cyclists. The whole data set and the supporting data are available open source for research purposes to ensure global accessibility.

Keywords Drone · Aerial observation · Interaction of different modes · Big data · Open data

Introduction

Common data sources for obtaining traffic information, such as loop detector data or trajectories from GNSS devices, are widely used in traffic engineering and management. However, these sources are either limited in their information content, e.g., only the vehicular flow and occupancy for single loop detectors, or they are restricted to certain modes of transport, such as bus or taxi GNSS data (Yu et al. 2023). Using video analysis from static cameras as well as equipped unmanned aerial vehicles has become a viable means for traffic monitoring (Barmounakis et al. 2016; Kim et al. 2019). With video monitoring, detailed trajectory data can be extracted for all road users in the observed camera field of view. This becomes more and more important to understand the behavior and the strategic, tactical, and operative

maneuvers of road users and to increase traffic safety and efficiency with this knowledge (Outay et al. 2020).

Trajectory data sets have been used for different types of analysis, proving its wide range of possibilities this kind of data provides. For example, it was investigated how speed behavior changes in the inner-city traffic network taking multimodality into account (Paipuri et al. 2021), microscopic parameters for driving behavior were analyzed (Pham et al. 2020) and lane choice and lane change behavior were considered (Espalader-Clapés et al. 2023a). Traffic engineering applications have also been investigated, for example, by estimating queue length profiles (Zhou et al. 2021). A comprehensive revisit on flow studies based on trajectory data can be found in Li et al. (2020).

Regarding safety analysis (Wang et al. 2019) and the environmental impact of traffic, detailed trajectory data can also give valuable insights. Based on a large-scale data set with vehicle trajectories of motorized road users (Barmounakis et al. 2021), the generated emissions and the impact on air pollution were considered. They developed an emissions-macroscopic fundamental diagram by including a

✉ Alexander Kutsch
alexander.kutsch@tum.de

¹ Chair of Traffic Engineering and Control, Technical University of Munich, Munich, Germany

microscopic emission model based on the trajectories. In addition, noise emissions in a congested environment were investigated and the correlation between noise emissions and traffic congestion was analyzed (Espalader-Clapès et al. 2023b).

These examples show that the need for data is becoming more relevant in the field of general traffic analysis and behavior, but also of individual traffic maneuvers like turning, overtaking, etc. It gets especially important when looking at the traffic behavior of vulnerable road users (VRU) as cyclists and pedestrians, as proper data sets about those road users are quite rare and therefore, knowledge about their behavior in the traffic system is quite limited.

For motorized traffic, most large-scale data sets also stem from observations on freeways, where the sensor density and the need for traffic detection are also higher, and therefore, data generation is easier. For urban traffic systems with a much higher variety of different road users and also a more complex road network structure, it is rather difficult to use (existing) infrastructure-based sensors like on freeways (Margreiter 2016; Motamedidehkordi et al. 2017) to have a full data coverage of all road users and their interactions. This present work goes one step further and elaborates on the use of drone video footage in combination with innovative recognition algorithms to provide a full temporal and spatial overview of the complete set of all road users. This enables further research to study in particular the interactions between different road users, especially motorized road users and VRUs with a focus on traffic safety and to better understand multimodal traffic patterns. Such data contributes significantly to improve traffic safety and to optimize the road network operations, as well as to train further video recognition algorithms and neural networks.

Existing Datasets

Mainly the first attempt for extracting trajectories from video data on a large scale is the NGSIM dataset (U.S. 2006). This dataset covers several hundred meters of traffic, mainly on major freeways, collected by several stationary cameras and only includes motorized traffic. Even though this data set was undoubtedly a pioneer and the basis for many findings, several studies have shown that the data set contains inaccuracies and anomalies (Punzo et al. 2011; Coifman and Li 2017). Further highway trajectories are publicly available for research purposes. The highD (Krajewski et al. 2018) and the exiD (Moers et al. 2022) dataset, published by the ika Aachen, provide information on free stretches, as well as exits and entries to German highways, respectively. Also collected by Unmanned Aerial Vehicles (UAV) on German highways, the AUTOMATUM DATASET (Spannaus et al. 2021) provides trajectory information initially

intended to validate automated driving functions. Shi et al. (2021) collected the High-Granularity Highway Simulation (HIGH-SIM) dataset using a helicopter, ensuring higher altitudes. This dataset also contains highly accurate highway trajectories.

Regarding urban traffic scenarios, mainly single intersections were taken into consideration up till now. Also published by the ika Aachen, the inD dataset (Bock et al. 2020) provides trajectories from four different urban intersections and the round dataset (Krajewski et al. 2020) for three roundabouts. For both data sets, the videos were collected using drones and they include pedestrians and bicyclists besides motorized traffic. The recordings are limited to the battery capacity, and therefore, several recordings exist for each location, but they are matched neither over time nor over space. The CitySim dataset (Zheng 2023) is a dataset that covers twelve different locations and over 1000 min of recording, also filmed by drones. Including a wide range of road geometries, both on intersections and freeway segments, it was initially supposed to contribute to safety research, as well as digital twins.

When it comes to covering a wider range of urban roads, the pNEUMA dataset (Barmounakis and Geroliminis 2020) is a seminal dataset including all modes of motorized traffic in the city center of Athens. Using a swarm of ten drones in parallel during the morning peak hours on five working days, around 1.4 km² was covered. Again, the recordings were limited due to battery constraints, resulting in recording lengths of around 15–20 min for each period. Therefore, the data are continuous in space but lack information on VRUs and the periods in between the flights. Recently, Yu et al. (2023) published a stationary camera-based dataset. For two cities in China, the authors used cameras at intersections distributed across the city to extract motorized traffic. They developed an algorithm to recognize the vehicles so that the trajectories could be tracked over the entire area and recognized across several cameras. Naturally, they could not directly track the unmonitored intermediate routes. A comparison of the mentioned data sets is shown in Table 1. A further overview of earlier data sets can be found in (Bock et al. 2020).

The limitations of the data sets listed above give rise to the need for (i) a coherent record covering several intersections, which is also (ii) continuous over time, (iii) depicts VRUs in addition to motorized traffic and (iv) is publicly available open source for research purposes. We close these gaps with the TUMDOT-MUC dataset (Trajectories from Urban Multimodal Drone Observations of Traffic—Munich), which is a trajectory dataset that was recorded using twelve drones during the afternoon peak hours on two weekdays. The experiment and the dataset are described in more detail in the following chapters. The data set was gathered within the Munich TEMPUS project (Kutsch et al. 2022), a project

Table 1 Comparison of different available trajectory data sets

Data set	Location	Available data	VRU	Time continuous	Space continuous	Collection device(s)
NGSim	Freeway, arterial segments	4 locations, 2.5 h	No	No	5–8 video cameras	Stationary mounted cameras
pNEUMA	Urban network	1.4 km ² , 5 week days, morning peak hours	No	No	10 drones	Swarm of 10 drones in parallel
highD	Freeway, segments	6 locations, 16.5 h	No	No	No	1 drone per location
exiD	Freeway, entries and exits	7 locations, 16 h	No	No	No	1 drone per location
inD	Urban intersections	4 locations, 10 h	Yes	No	No	1 drone per location
roundD	Roundabouts	3 locations, 6 h	Yes	No	No	1 drone per location
HIGH-SIM	Freeway	1 location (2400 m), 2 h	No	No	3 video cameras	helicopter with 3 cameras
CitySIM	Freeway, expressway, intersections	12 locations, 19 h	No	No	1–2 drones	1 or 2 drone(s) per location
AUTOMATUM—Autobahn Open	Freeway segments	Several locations, 30 h	No	No	No	1 drone per location
City-scale vehicle trajectory data	Urban intersections with reconstructed trajectories	4 days, 12–24 h per day, 441–1838 cameras	No	Yes	Yes, with reconstructions	Fixed cameras

Fig. 1 The six observation locations in the North of Munich (Own elaborations based on Google Earth Pro)

to prepare the Munich urban and suburban road network for connected and automated driving.

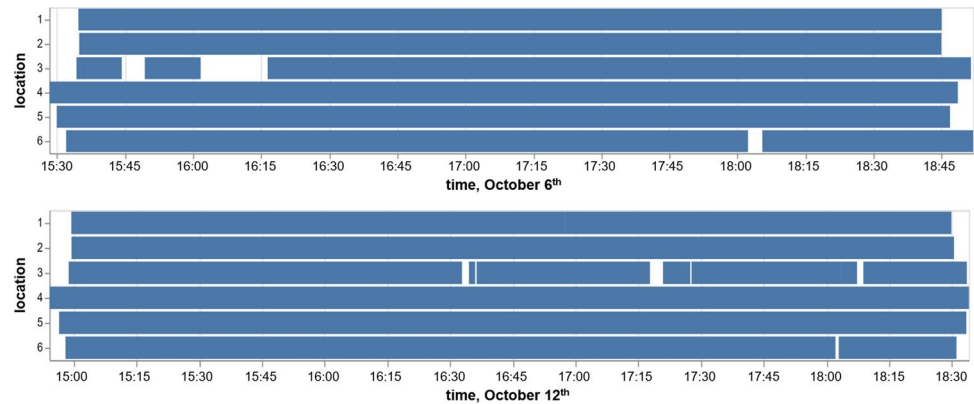
Experimental Design

The drone surveys were conducted along Rheinstrasse in Munich, Germany, between Bonner Platz and Leopoldstrasse. Six locations were covered by two drones each, as shown in Fig. 1. The drones had a flying altitude of 110 m when taking the videos, resulting in a recording area of 143 m on 75 m for each drone location. With the need for an overlap between the individual locations for reconstructing the complete trajectories throughout the whole stretch, a total length of 700 m was recorded. The road is highly frequented and can be classified as a main road. The perimeter includes two signalized intersections, one right-of-way-controlled intersection, two non-signalized T-intersections

and two signalized pedestrian crossings. Bicycle traffic in this area is mostly guided on a protective lane. For the general traffic, and pedestrians in particular, there are several points of interest within the recording area and the immediate surroundings. In addition to a school and a kindergarten, several doctors and pharmacies, as well as numerous shopping facilities, both for groceries and drugstores, and recreational activities are also available. Several restaurants, a sports club and three hotels are located in the area. There are also several access points to public transport, namely a tram and several bus stops on Leopoldstrasse (location 6), as well as a subway access point on Bonner Platz (location 1).

Until today, drones have, assuming optimal conditions, a maximum flight time of around 30 min due their limited battery capacity. Various systems, such as balloons and tethered drones with power supply, have been assessed but do not yet provide the necessary image stability or are not suitable due to the high costs and the more complicated approval process,

Fig. 2 The observation time-lines for both days, including recording gaps



especially in the city center. For this study, in order to create a continuous data set, two drones were used for each location, and a ‘shake-hands’ was performed in the air so that the drone to be replaced only descended when the following drone was already in the air and had taken over and fixed the image section. Accordingly, a total of twelve drones were in use. They recorded a total of analyzed videos amounting to almost 42 h from their respective single locations. The recordings were performed on Thursday, October 6, 2022 and Wednesday, October 12, 2022. The days in the middle of the week were determined according to surrounding detector data with the purpose to capture the most frequented times in the week. On both days, flights were performed in the afternoon in order to be able to depict the leisure traffic as well as the commuter traffic and therefore the peak-hour. On the first day, the last drone was in the air at 15:35, so that a complete image of the entire stretch could be generated. The recordings were carried out until 18:45. On the second day, the first shots were started even earlier to counteract the earlier sunset. Thus, from 15:00 on all drones were in the air until the end of the recording at 18:25.

Due to their altitude and a mostly unobstructed viewing angle toward the observed road and sidewalk segments, the drones were able to record all road users in the designated area. The different classified modes are passenger cars, buses, trucks, trams, motorcycles, (e-)bicycles, pedestrians and (e-)kick-scooters. Moreover, truck trailers and bus trailers (in the city of Munich several high-capacity bus routes operate busses with bus trailers) were also recognized and categorized. Please note that no distinction could be made between ordinary bicycles, electric bicycles and cargo bicycles as these could not be differentiated from the corresponding flight altitude. In addition to those road users, the dataset also contains trajectories of an electrified bicycle Rickshaw (Fehn et al. 2023; Margreiter et al. 2023) as well as the trajectories of two semi-automated vehicles as part of the TEMPUS project.

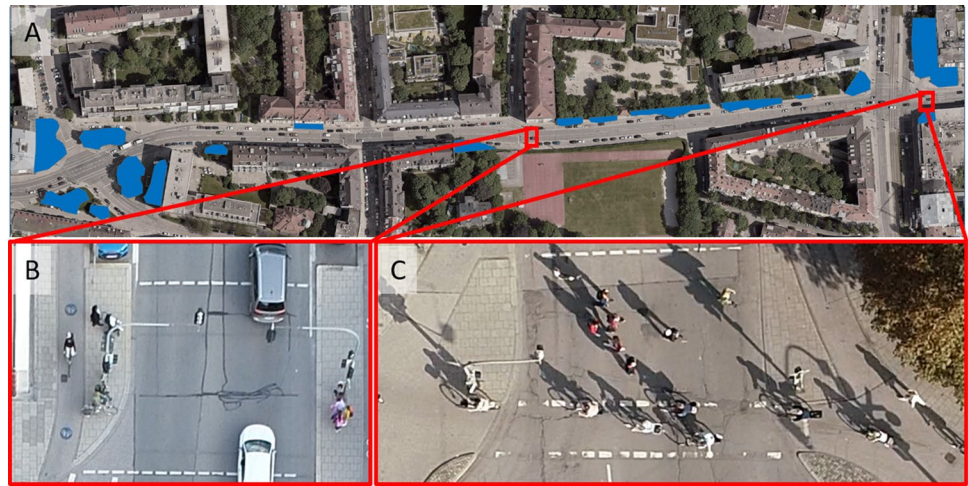
Due to the high complexity of coordinating twelve drones, isolated technical problems (activation of the drone’s return

home function when the battery was almost fully charged, inexplicable distortion of the videos), as well as strong rotations of individual drones due to wind gusts, there were some isolated monitoring gaps during the recordings, which are shown in Fig. 2. This mainly affected location 3 and, for two short periods, location 6. At the other locations, the monitoring and the changeover in the air proceeded smoothly.

The continuous temporal coverage requires additional effort in terms of planning and resources. As far as the recording area is concerned, twice as many drones, pilots and batteries are required to cover the same area. Alternatively, all drones, in this case twelve, could be used in parallel. This would result in twice the area covered. Finally, the resources could also cover a more extended period than just the planned peak hours. In all cases, there would be a time gap when the battery is changed and valuable information would be lost. However, the following aspects supported continuous recording for this experiment:

1. The area chosen for the survey was selected, so that there is a high volume of traffic and covers many pedestrians and cyclists. A local extension would still be valuable but not that much related to the chosen focus as the surrounding roads do not have a similar impact on VRU.
2. The effective recording time per drone was between 15 and 18 min. Assuming 10 min of idle time for technical tasks and coordination, the number of trajectories would amount to 55–65% compared to full observation. If the observation time were extended, i.e., by adding further recordings before and after the chosen recording hours, it would be possible to achieve the same amount of trajectories. However, the loop count values of the adjacent detectors show that traffic decreases significantly outside the peak hours. For the detector at the eastbound approach to Leopoldstrasse, the traffic volume for the peak hours between 15:00 and 18:00 on the second day of recording was around 385 vehicles/hour. In contrast, the traffic counts between 12:00 and 15:00 are roughly 270 vehicles/hour. Likewise, the number drops

Fig. 3 Visual obstructions in the recording area due to trees and construction sites shown in blue (A) (Own elaborations based on Google Earth Pro) and visualization of problematic scenarios due to static objects, such as traffic signals (B) and shadows (C) in crowded areas (zoomed view) (Colour figure online)



from 290 vehicles/hour at 18:00 to 170 vehicles/hour at 20:00. Therefore, the trajectories would be 25–55% less in these periods, and almost double the time would be needed to cover the same amount of trajectories.

3. Traffic engineering-related aspects, such as the traffic state and the formation of waiting queues, are very time-dependent and related to the near past states. Accordingly, a continuous data set is valuable for estimation approaches that include the short-term past.

Processing Steps

Besides the missing time gaps, further implications arose from shadowing caused by trees and construction sites due to the bird's eye view of the drones. These visual obstructions are shown in blue in Fig. 3A. The obstructions only affect the pathways for pedestrians and bicyclists and mainly occur at location 1 and on the northern sidewalk of Rheinstrasse. In addition, in some places, it is difficult even for the human eye to distinguish whether an object is a dynamic object, a static object or a shadow, despite the comparatively low flight altitude. Figure 3B shows the signalized crossing from location 4. Here, for example, the light signal post is often misclassified as a person and the distinction between individuals in the waiting area is regularly blurred. Figure 3C shows the eastern crossing of Potsdamer Strasse at the signal ford. Here, too, the large number of people and the low sun make it difficult to distinguish between the individual road users and their shadows. This problem also occurs with two people walking very close to each other. Several filters and processing steps were performed to remove misclassifications, connect trajectories separated by visual obstructions and sort out static objects.

The raw data, as provided by DeepScenario GmbH, consists of one data file per evaluated frame in each individual

video, i.e., for each location and drone flight, and a matching file of the trajectories for each timeline on one location. To obtain these timelines for each location, the consecutive videos were connected at one location using manual frame matching, i.e., the frame of the ending video was compared with the corresponding frame of the following video, provided there were no gaps in the recording. The evaluation frequency was set to 12.5 frames per second, i.e., providing updated information every 0.08 s. The extracted trajectories contain information on the track (or object) ID, classification, coordinates, dimensions, speed and acceleration of each tracked object in each frame. Based on this information, the following steps were implemented to get a clean and spatially and temporally continuous dataset:

(i) Static objects, which do not move throughout the whole observed period, were filtered out. These trajectories belong to misclassified objects, such as traffic light poles, postal mailboxes or parked vehicles. This filter is based on a defined displacement threshold, which was chosen to be two meters in our case. The center point of the object during the tracked period was first determined on the 2D space, i.e., the longitude (x) and latitude (y) coordinates. The Euclidean distance to this center point was then calculated for each point of the trajectory in time. If the amount of timestamps, where the distance is below the threshold value, exceeds 80%, the trajectory was removed.

(ii) For non-motorized road users, a second filter was used to connect different trajectories that belong to the same object but are not assigned to the same ID due to obstruction or other reasons. This is summarized in the following steps:

1. Category, position and speed at the first and last point in time of the trajectory were determined for each object. As the speeds of objects are specified in terms of direction as x - and y -vectors according to the axis of the global coordinate system (the third dimension, i.e., the

z-vector, was not used for filtering), the average of the values for the last second was used for the final speed in order to include changes in direction.

2. For each object, it was checked whether there is an object with a higher ID (i.e., an object that was detected later than the currently considered one), where (a) the time gap between the last detection of the current object and the starting point of the following object is below a defined threshold time gap and (b) whose category matches. From testing different threshold time gaps, a maximum of 250 frames turned out to perform best for all VRUs.
3. A calculated position of the ending trajectory was then computed for the determined time gap using the end speed vectors described before and the actual time difference.
4. If the distance between the calculated position from step 3 and the start position of the following object is less than a defined threshold distance, both trajectories are assigned the same ID. The threshold distance was set at 4.0 m.
5. The gap values that could not be recorded directly were then filled by linear interpolation. Interpolated values are marked in the data set with an additional attribute to prevent the illusion of unmonitored ground truth, for example, for behavioral models.

(iii) The data for the individual locations were then once again filtered for trajectories existing at least 0.5 s.

(iv) As a last step concerning the individual videos, a manual check was carried out in order to eliminate further errors as much as possible.

(v) The resulting data, which is accordingly available isolated for the individual locations and time slots, then had to be linked in terms of time and space. Due to small localization and dimension errors of up to 15–20 cm, isolated misclassifications, errors in the time synchronization of the drones, as well as small deviations in the detected speeds, the linking of the trajectories is not quite trivial, especially for pedestrians and cyclists.

A synchronization file connecting the ending frame of one video to the corresponding start frame of the following video was manually created for continuously matching the consecutive videos on one location. For each video transition in time, a matching score S was calculated for all objects appearing in the ending frame and the starting frame of the following video, wherever the spatial distance was less than five meters in between. For this purpose, each object's coordinate, labeled x and y , their speed vectors v_x and v_y , referenced in the global coordinate system, as well as their dimensions in the objects coordinate system, i.e., the width d_x and the length d_y were utilized. This score, shown in formula 1, was then used to assign the

corresponding object IDs. A value close to 0 indicates the same object, as the current position, the speed and the area covered by the objects are nearly equal. A value higher than 3 turned out not to be a match in any case assuming the correct transition frame.

$$S = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} + \sqrt{(v_{x_2} - v_{x_1})^2 + (v_{y_2} - v_{y_1})^2} + |(d_{x_1} * d_{y_1}) - (d_{x_2} * d_{y_2})| \quad (1)$$

From the previous filters, objects tracked in one video might not appear in the consecutive video. For example, a car that performs a parking maneuver in a video is present in the corresponding data. However, if this car is parked for the duration of the subsequent video, it will be filtered out there. Consequently, objects without a match in the respective other data file might eventually appear. These cases retain with their existing ID. In cases where more than one pair of objects had a score below 3, the pair with the lowest score was matched. The IDs of the merged objects were adjusted so that the ID from the first video was retained. The remaining objects received a new, globally unique ID and the timestamps were converted from individual videos to a global timeline.

(vi) The continuous timelines were then merged across the six locations. As described above, the data set at this point contained a unique ID for each object in one timeline per location and day. To track individual trajectories throughout the whole observation area, the respective track ID needed to be matched between the locations as well. The linking logic was implemented according to the following description to guarantee a scalable approach.

Before starting to concatenate, a reference location needed to be determined. To ensure the correct time for each location and since the start-times vary slightly between the drones, location 2 was chosen as reference timeline for this. This means that the first frame of the first video of location 2 is set to be the global frame 0 as it provides a continuous recording throughout both days. Starting from this location, all other locations were added one by one and the global time stamps were adjusted. The process for merging locations 2 and 3 is explained below as an example.

1. The overlapping areas of the first videos of both locations were manually compared to find the rough time difference (to the second) between the beginnings of the videos, as shown in Fig. 4. For location 3 on the first day, the second 26 (i.e., roughly 26 s with 25 frames per second = frame 650) equals the start of location 2.

Fig. 4 Overlapping areas in the respective first videos of location 2 and location 3 on the first recording day. An offset of roughly 26 s was found and used as input for merging the trajectories of these locations

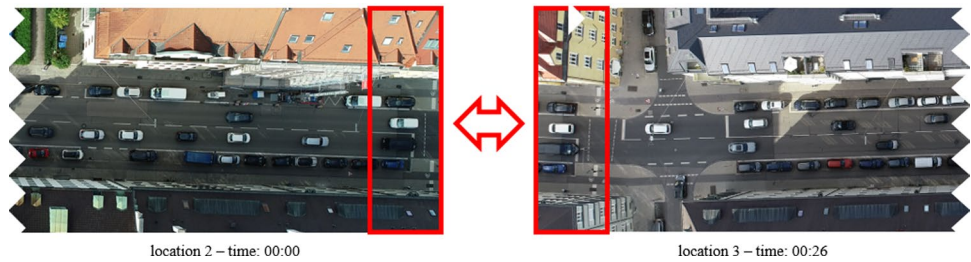
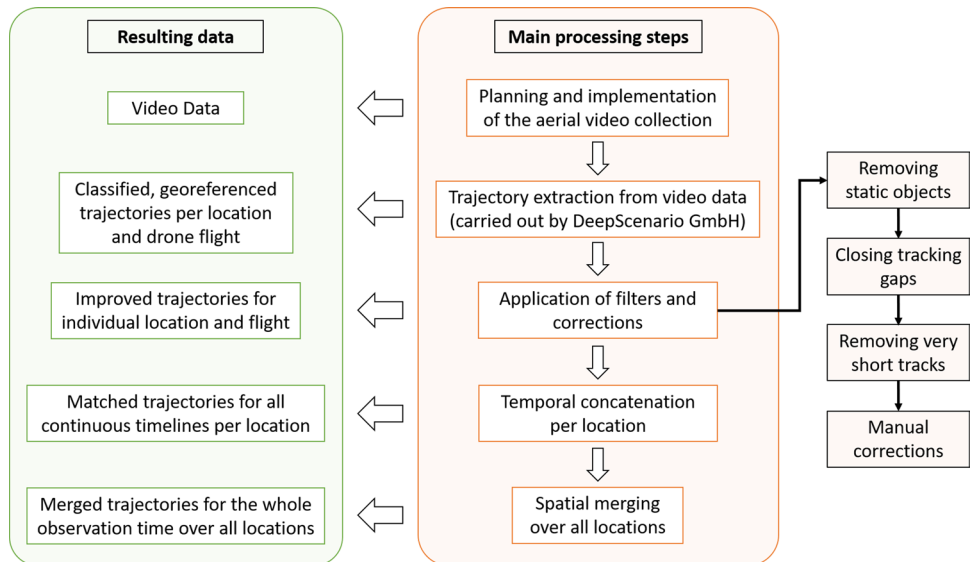


Fig. 5 Pipeline for recording and processing the data in order to obtain a time and space-continuous image of the traffic



- Based on this difference, the previously used score S , described in Eq. 1, was determined for all objects for the period of 100 frames before and after this determined matched frame in the filtered dataset. Again, only scores below three were kept, as well as the lowest score for duplicate matches was used.
- The frame in the search area with the highest number of matches and the lowest sum of all scores was retained as the matching frame. Therefore, frame zero from location 2 was compared with every frame between 550 and 750 from location 3 – and the frame with the best score was stored.
- This process was carried out for the first 100 s (i.e., the first 1250 frames) of each timeline, and again, for each comparison, the frame with the best score and maximum number of matches was stored. The time difference that occurred most frequently with the method described was defined as the actual time difference.
- This frame difference was then used to assign the objects in the overlapping area for each frame in the timelines of the neighboring locations. The matched IDs were then again assigned to the same object ID, whereas the non-matched objects received a new globally unique ID.
- In the overlapping area, only the entries of location 2 were kept to eliminate duplicate data points. Finally, the timestamps of location 3 were adjusted to match those from location 2.

This process was repeated for each timeline after an observation gap as shown in Fig. 2, resulting in a complete merge of location 2 and location 3 on one day. As soon as the two locations were completely processed, the next location was added to the data in the same way. As a result, the final data sets with connected trajectories and global timestamps were retrieved. The whole processing pipeline is visualized in Fig. 5. The resulting merged data set, which is available open source, is explained in more detail in the following section.

Format Description

The processed data, split into 10 min sections for easier handling, are available in csv-formatted files to ensure easy access. Each file contains the trajectories of all road users in the entire survey area in the corresponding period. One

Table 2 Extracted data parameters from the drone video footage

Parameter	Description	Example value	Unit
Timestamp	Numeric global timestamp	105.12	[s]
Category	Numeric classification ID	1	[–]
Track_id	Numeric unique ID	38	[–]
Translation	[x, y, z]	[123.0319, 14.0211, 508.6254]	[m]
Dimension	[length, width, height]	[4.2998, 1.9754, 1.4882]	[m]
Rotation	[x, y, z]	[– 0.0011, 0.0016, 0.1227]	[–]
Velocity	[x, y, z]	[10.252, 1.2638, – 0.0169]	[m/s]
Acceleration	[x, y, z]	[0.0577, 0.0071, – 0.0001]	[m/s ²]
Interpolated	Numeric value	0	[–]

line describes an object at a point in time with the attributes shown in Table 2.

The timestamp describes the global time in seconds, starting from the very first reference frame. This is for both days the first frame the recording started at each location, i.e., when all drones started recording in parallel. The object category is stored as a numeric value for each category mentioned before (truck, bicycle, pedestrian, etc.). Category ID 1, as shown in the example, would match the classification as a passenger car. The track ID is the object identifier for each road user. This value is unique for the whole data sets, making it possible to follow objects throughout the different subsequent frames. The translation describes the location of the respective object, where the values describe the offset with respect to a reference point, which is given in the UTM 32 N (NIMA 1989) coordinate system and is uniform for the entire data set. The reference point and the assignment table for the object classifications, as well as other meta information, are available together with the data set. The translation value refers to the center of the bounding box and the bottom of the object, i.e., on the ground. As in the global coordinate system, one unit equals one meter. In addition to the position, the three-dimensional bounding box, i.e., the length, width and height of the object are also shown in meters. The current velocity and the acceleration are given as a vector in the respective global coordinate system and the rotation describes the current orientation in that system. The rotation vector refers to the respective rotational axis, i.e., the z value refers to the orientation parallel to the surface of the earth. It can easily be transformed into any other representation, for example with the `scipy.spatial.transform.Rotation` module.¹ As the processing steps include interpolation steps for the obstructed areas, each data point is marked as 1 if it is interpolated and 0 if it comes directly from the computer vision algorithm supplied by DeepScenario GmbH.

Additional Data

In addition to the trajectory data itself, supporting material is also available. The loop detectors shown in red in Fig. 6 provide counts of flow in vehicles per interval, aggregated for 15 min intervals, as well as occupancy values between 0 and 1. The loop detectors are inductive single loops. Therefore, no velocities or classifications are detected. One detector per lane is located in front of each intersection approach. The values are provided for the periods of the surveys on both recording days and can be downloaded together with the trajectories.

Furthermore, the switching data of the traffic signal systems shown in yellow in Fig. 6, which contain the actual signals on a second-by-second basis, are also supplied. Each signal group is assigned a unique ID. The data table contains the attributes signal ID, a timestamp and the displayed signal. One line corresponds to the display status of one signal group. A signal change is saved in a new row with the same signal ID, the timestamp to the second and the updated signal. Consequently, the difference between the two timestamps corresponds to the display duration of the first signal.

In order to be able to match the trajectories to a map, an OpenDRIVE map as specified by the Association for Standardization of Automation and Measuring Systems (ASAM)² is supplied, containing the full ground truth of the street layout and the infrastructure of the recording area.

Resulting Dataset and Limitations

The resulting data set contains over 24,000 trajectories per day. A more detailed look at the trajectories on the second day of the survey (October 12) shows that out of a total of 24.138 trajectories, 10.531 originate from motorized road users, i.e., cars, motorcycles, busses, heavy goods vehicles or streetcars (i.e., trams), see Fig. 7A. Consequently, around

¹ Documentation available at <https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.transform.Rotation.html>

² Detailed description available at <https://www.asam.net/standards/detail/opendrive>

Fig. 6 Single inductive loop detectors (red) and traffic signals (yellow) of which the data can be downloaded additionally. (Own elaborations based on Google Earth Pro) (Colour figure online)

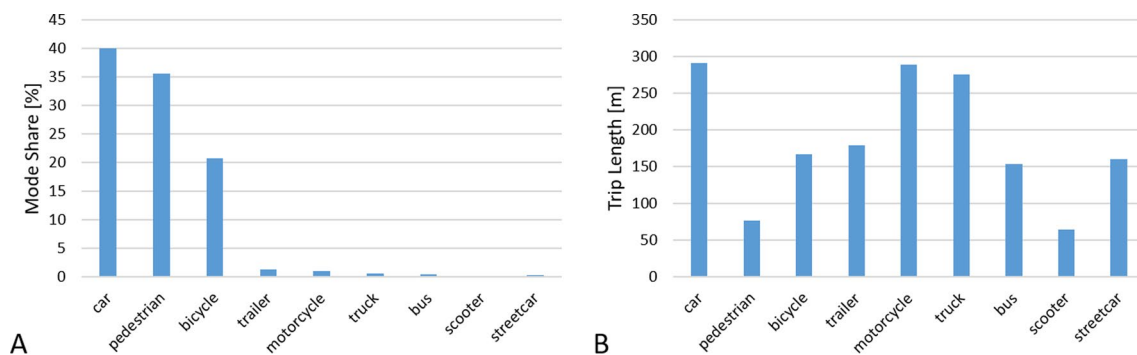
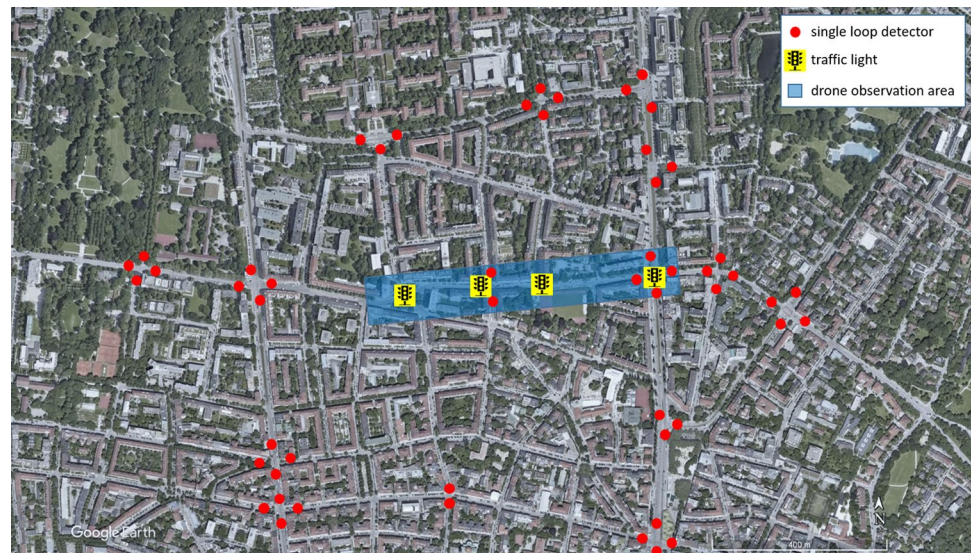


Fig. 7 The mode share of the collected trajectories (A) and the average trip length per category (B)

56% represent VRUs, i.e., pedestrians, cyclists and scooters. The average trajectory length is 186 m, where motorized road users appear to have longer trips, as can be seen in Fig. 7B.

At this point, it should be mentioned again that there were problems with obstructions from trees and considerable difficulties with people walking in groups and being very close together. These errors could not completely be corrected in post-processing, which is why trajectory breaks and abrupt tracking changes between two people occur. The actual number of VRUs is, therefore, smaller than the number of trajectories, and the average trip length is somewhat longer. These errors do not exist for motorized traffic, which, on the one hand, uses the uncovered road and, on the other hand, is sufficiently large to always differentiate clearly.

The data can be used to quickly analyze traffic patterns. By having all trajectories available, capacity bottlenecks can be easily identified. Figure 8 shows the average speed of all motorized road users in the period from 16:00 to 16:15. The area was divided into cells of equal size, each with an edge

length of 0.25 m, and each data point was assigned using the position corresponding to the cell. An average speed was then determined for each cell from all the data points and displayed. Blue cells represent a slow average speed, orange and yellow area represent high up to free flow speed. In particular, it can be seen that the eastbound Rheinstrasse in the approach to Leopoldstrasse has significantly lower average speeds than the opposite direction. The illustration realistically depicts the problem of long queues and major time losses.

Conclusions and Outlook

The highly detailed dataset allows applications in traffic flow theory to be developed, validated, and calibrated. For example, the immense ground truth data set can be used to develop and verify car following models, turning behavior models, link-level traffic flow models, or queue length estimators. In addition, the multimodal trajectories, which

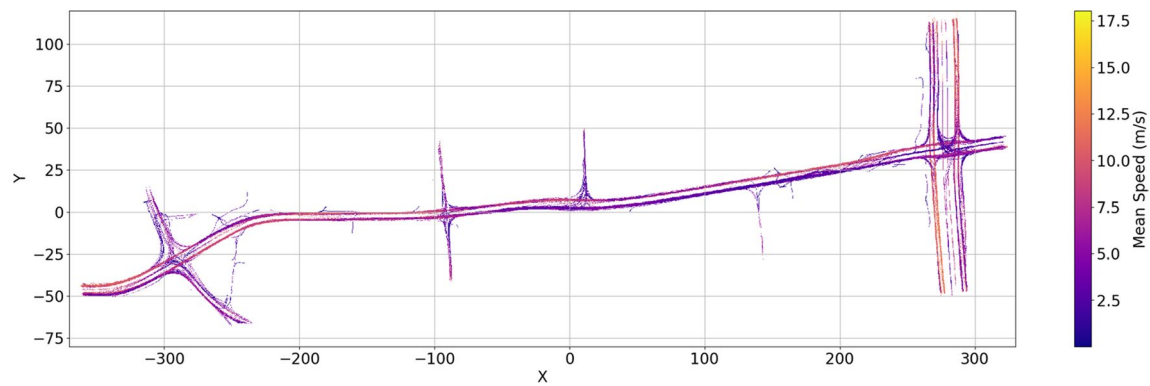


Fig. 8 Mean speed in the observation area between 16:00 and 16:15 on the second recording day

include non-motorized road users with a high level of detail, also depict interactions between VRUs and motorized vehicles, both on the open, non-disturbed link, as well as at signalized and non-signalized intersections. Consequently, especially in the presence of different levels of vehicles' connectivity and automation, important conclusions can be derived regarding the impact on traffic flow in urban environments.

Also, regarding traffic safety aspects, having each road user together with the information on the underlying road network, different situations have been captured. Having several different types of intersections, as well as different types of bicycle paths and other infrastructure elements, the dataset is ideally suited for a more detailed safety analysis.

To enable a wide and global access for scientists and practitioners, the TUMDOT-MUC data set is available open source. The open data set includes all extracted trajectories for all road users continuously in time and space through the whole area covered by all twelve drones at the six locations, as well all the supporting data sources mentioned before. It is available under the Creative Commons CC BY-NC 4.0 license to guarantee easy access and usability for further research.

In beginning of summer 2023, a similar dataset was gathered in the city of Ingolstadt, Germany, via the KIVI project (Ilic et al. 2022a, 2022b) funded by the German Federal Ministry of Digital and Transport. The number of drones was increased to a total of 18, which flew at a reduced altitude of 100 m. This enabled an even greater accuracy to be achieved and a larger area to be mapped, in total 9 locations. Also for this study in Ingolstadt, the afternoon peak hours were chosen for the observation time, with recordings being made on a total of three days. Likewise, this data set, TUMDOT-IN, will be made available open source.

Acknowledgements The drone data were gathered with the help of Fairfleet GmbH (Drone Flights) and DeepScenario GmbH (Trajectory Generation).

Author Contributions Conceptualization: AK, MM, KB; data curation: AK; formal analysis: AK; funding acquisition: MM, KB; investigation: AK, MM; methodology: AK, MM; project administration: MM; supervision: KB; writing – original draft: AK, MM; writing – review and editing: MM, KB.

Funding Open Access funding enabled and organized by Projekt DEAL. This research was funded and supported via the TEMPUS project through the German Federal Ministry for Digital and Transport under the research grant 01MM200008K.

Data Availability The data set can be downloaded for research purposes at <https://www.mos.ed.tum.de/en/vt/research/data-sets/tumdot-muc/>.

Declarations

Conflict of Interest The authors do not have any financial or non-financial conflicts of interest to declare.

Ethical Approval The recordings were conducted in consultation with the relevant authorities. Due to the flight altitude, no personal data is recognizable. Therefore, no explicit consent was obtained from road users.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Barmounakis E, Geroliminis N (2020) On the new era of urban traffic monitoring with massive drone data: the pNEUMA large-scale field experiment. *Transp Res Part c Emerg Technol* 111:50–71. <https://doi.org/10.1016/j.trc.2019.11.023>

- Barmounakis E, Vlahogianni E, Golias JC (2016) Unmanned aerial aircraft systems for transportation engineering: current practice and future challenges. *Int J Transp Sci Technol*. <https://doi.org/10.1016/j.ijtst.2017.02.001>
- Barmounakis E, Montesinos-Ferrer M, Gonzales EJ, Geroliminis N (2021) Empirical investigation of the emission-macroscopic fundamental diagram. *Transp Res Part d Transp Environ*. <https://doi.org/10.1016/j.trd.2021.103090>
- Bock J, Krajewski R, Moers T, Runde S, Vater L, Eckstein L (2020) The inD Dataset: A Drone Dataset of Naturalistic Road User Trajectories at German Intersections. 2020 IEEE Intelligent Vehicles Symposium (IV), Las Vegas, NV, USA. <https://doi.org/10.1109/IV47402.2020.9304839>
- Cofman B, Li L (2017) A critical evaluation of the next generation simulation (NGSIM) vehicle trajectory dataset. *Transp Res Part b Methodol*. <https://doi.org/10.1016/j.trb.2017.09.018>
- Espadaler-Clapés E, Barmounakis E, Geroliminis N (2023a) Empirical investigation of lane usage, lane changing and lane choice phenomena in a multimodal urban arterial. *Transp Res Part a Policy Pract*. <https://doi.org/10.1016/j.tra.2023.103674>
- Espadaler-Clapés J, Barmounakis E, Geroliminis N (2023b) Traffic congestion and noise emissions with detailed vehicle trajectories from UAVs. *Transp Res Part d Transp Environ*. <https://doi.org/10.1016/j.trd.2023.103822>
- Fehn F, Engelhardt R, Margreiter M, Bogenberger K (2023) Ride-Parcel-Pooling: Integrating On-Demand Passenger Transportation and City Logistics. PIARC World Road Congress, Prague, Czech Republic.
- Ilic M, Sautter N, Margreiter M, Loder A, Bogenberger K (2022) Changing Strategic Alignments in European Urban Traffic Control - Requirements for Future Developments. 10th Symposium of the European Association for Research in Transportation (hEART), Leuven, Belgium.
- Ilic M, Sautter N, Margreiter M, Loder A, Bogenberger K (2022) Revisiting Urban Traffic Control in Austria, Germany and Switzerland - Requirements for Future Developments. Transportation Research Board - 101st Annual Meeting Washington, D.C USA.
- Kim EJ, Park HC, Ham SW, Kho SY, Kim DK (2019) Extracting vehicle trajectories using unmanned aerial vehicles in congested traffic conditions. *Mach Learn Transp*. <https://doi.org/10.1155/2019/9060797>
- Krajewski R, Bock J, Kloeker L, Eckstein L (2018) The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems. 21st International Conference on Intelligent Transportation Systems (ITSC), <https://doi.org/10.1109/ITSC.2018.8569552>.
- Krajewski R, Moers T, Bock J, Vater L, Eckstein L (2020) The rounD Dataset: A Drone Dataset of Road User Trajectories at Roundabouts in Germany. IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), Rhodes, Greece. <https://doi.org/10.1109/ITSC45102.2020.9294728>.
- Kutsch A, Margreiter M, Stüger P, Fehn F, Metzger B, Hamm L, Engelhardt R, Ilic M, Bogenberger K (2022) TEMPUS - Test Field Munich - Pilot Test for Urban Automated Road Traffic. IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), <https://doi.org/10.1109/ITSC55140.2022.9922303>.
- Li L, Jiang R, He Z, Chen X, Zhou X (2020) Trajectory data-based traffic flow studies: a revisit. *Transp Res Part c Emerg Technol*. <https://doi.org/10.1016/j.trc.2020.02.016>
- Margreiter M, Ilic M, Joshi A, Bogenberger K (2023) A Self-Driving Cycle-Rickshaw for Autonomous Urban Passenger and Freight Transport, PIARC World Road Congress, Prague, Czech Republic.
- Margreiter M (2016) Fast and Reliable Determination of the Traffic State Using Bluetooth Detection on German Freeways. World Conference on Transport Research (WCTRS), Shanghai, China.
- Moers T, Vater L, Krajewski R, Bock J, Zlocki A, Eckstein L (2022) The exiD Dataset: A Real-World Trajectory Dataset of Highly Interactive Highway Scenarios in Germany. IEEE Intelligent Vehicles Symposium (IV), Aachen, Germany. <https://doi.org/10.1109/IV51971.2022.9827305>.
- Motamedidehkordi N, Margreiter M, Hoffmann S, Busch F (2017) Implications of Automated Vehicles on Freeway Safety and Operations, ITS World Congress Montreal, Canada.
- NIMA - US National Imagery and Mapping Agency (1989) The Universal Grids: Universal Transverse Mercator (UTM) and Universal Polar Stereographic (UPS). Edition 1, TM 8358.2 Washington, D.C., USA.
- Outay F, Mengash HA, Adnan M (2020) Applications of unmanned aerial vehicle (UAV) in road safety, traffic and highway infrastructure management: recent advances and challenges. *Transp Res Part a Policy Pract*. <https://doi.org/10.1016/j.tra.2020.09.018>
- Paipuri M, Barmounakis E, Geroliminis N, Leclercq L (2021) Empirical observations of multi-modal network-level models: Insights from the pNEUMA experiment. *Transp Res Part c Emerg Technol*. <https://doi.org/10.1016/j.trc.2021.103300>
- Pham HQ, Camey M, Pham KD, Pham KV, Rilett LR (2020) Review of Unmanned Aerial Vehicles (UAVs) Operation and Data Collection for Driving Behavior Analysis. CIGOS 2019, Innovation for Sustainable Infrastructure. https://doi.org/10.1007/978-981-15-0802-8_178.
- Punzo V, Borzacchiello MT, Ciuffo B (2011) On the assessment of vehicle trajectory data accuracy and application to the next generation simulation (NGSIM) program data. *Transp Res Part c Emerg Technol*. <https://doi.org/10.1016/j.trc.2010.12.007>
- Shi X, Zhao D, Yao H, Li X, Hale DK, Ghiasi A (2021) Video-based trajectory extraction with deep learning for high-granularity highway simulation (HIGH-SIM). *Commun Transp Res*. <https://doi.org/10.1016/j.commtr.2021.100014>
- Spannaus P, Zechel P, Lenz K (2021) AUTOMATUM DATA: Drone-based highway dataset for the development and validation of automated driving software for research and commercial applications, IEEE Intelligent Vehicles Symposium (IV), Nagoya, Japan. <https://doi.org/10.1109/IV48863.2021.9575442>.
- U.S. Federal Highway Administration (2006) Next generation simulation (ngsim). <https://ops.fhwa.dot.gov/trafficanalysisitools/ngsim.htm>.
- Wang C, Xu C, Dai Y (2019) A crash prediction method based on bivariate extreme value theory and video-based vehicle trajectory data. *Accid Anal Prev*. <https://doi.org/10.1016/j.aap.2018.12.013>
- Yu F, Yan H, Chen R et al (2023) City-scale vehicle trajectory data from traffic camera videos. *Sci Data* 10:711. <https://doi.org/10.1038/s41597-023-02589-y>
- Zheng O, Abdel-Aty M, Yue L, Abdelraouf A, Wang Z, Mahmoud N (2023) CitySim: a drone-based vehicle trajectory dataset for safety-oriented research and digital twins. *Transp Res Rec*. <https://doi.org/10.1177/03611981231185768>
- Zhou Q, Mohammadi R, Zhao W, Zhang K, Zhang L, Wang Y, Roncoli C, Hu S (2021) Queue Profile Identification at Signalized Intersections with High-Resolution Data from Drones. 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), Heraklion, Greece. <https://doi.org/10.1109/MT-ITS49943.2021.9529337>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.