Spatial analysis and modelling of bicycle accidents and safety threats

Martin Loidl*, Robin Wendel#, Bernhard Zagel†

* # † Department of Geoinformatics, GI Mobility Lab
University of Salzburg
Hellbrunnerstraße 34, A-5020 Salzburg, Austria
e-mail: [martin.loidl;robin.wendel;bernhard.zagel]@sbg.ac.at

ABSTRACT

In this paper geographical information systems (GIS) are employed to analyse and better understand bicycle accidents as spatial incidents. Gained insights are then further processed in integrated, spatial models which in turn are building blocks of planning and information tools that explicitly address the issue of bicycle safety. Based on two use cases, this paper aims to demonstrate the additional value that can be generated by an integrated, explicitly geospatial approach. Existing methods and tools from the transportation and geographical community are merged and enriched with geospatial components.

Keywords: geospatial analysis, geospatial modelling, planning.

1 INTRODUCTION

In order to meet the goals of national and transnational safety initiatives, such as vision zero [1], it is of great importance to improve the situation of vulnerable road users, particularly of cyclists. Apart from this, safety concerns are a major factor that keeps people from cycling, although the benefits in terms of health effects and sustainability are obvious [2]. Therefore the provision of safe infrastructure and the addressing of relevant concerns among citizens must be an integrated part of any bicycle promotion strategy. Both aims can be achieved when safety aspects are explicitly considered in the analysis of existing and planning of new road infrastructure as well as in user-tailored information applications.

Transportation systems are manifested in the physical space and can thus be modelled and analysed from a geographical perspective. Geographical information systems (GIS) allow for an integrated, space- and time-sensitive perspective on transportation systems and mobility. Thereby, the geographical coordinates ("geospace") serve as common denominator for all transportation-related information, from infrastructure to moving objects and singular events. Following geographical concepts [3], GIS can be utilized in numerous ways for a better understanding of bicycle safety as phenomenon in the spatial and temporal dimension. Results from geospatial models and analyses are then further employed for evidence-based planning tools and decision support systems [4].

In the following we argue that bicycle accidents are spatial phenomena by their very nature. Thus spatial analysis approaches can reveal where and under which conditions bicycle accidents occur [5]. After a brief review of the status quo, two use cases will demonstrate the benefit of an explicit geospatial approach. In use case I existing studies will be extended by an investigation of spatial (and temporal) characteristics of bicycle accidents on a single incident level. Use case II deals with prototypical application examples for the utilization of geospatial information in planning and information tools for safer bicycling.
2 STATUS-QUO AND MOTIVATION

Transportation geography and the utilization of geographical information systems (GIS) in the transportation domain (GIS-T) have a long tradition. Transportation is about movement in space and thus geographical concepts are relevant for modelling, analysing and finally better understanding of transportation systems. Due to the growing capabilities and increasing processing power, geographical information systems have evolved from static model frameworks, analysis tools and visual interfaces [6] to highly dynamic environments which can be utilized for numerous transportation issues [7].

Although the spatial nature of transportation, where bicycle traffic can be regarded as a subgroup, is obvious, the implications are often neglected. One of the reasons for this is indicated by Shaw [7]: GIS experts and transportation experts may work on the same topic, but with diverse “mind-sets”. Thus a cross-domain communication of experts from various disciplines is required as it can be expected to leverage yet isolated approaches. With regard to bicycle safety the author has demonstrated in a recent review article [8] the benefits of an explicit spatial perspective, based on various application examples.

Nevertheless it can be stated, that potential synergies of a tighter collaboration of experts dealing with bicycle traffic and safety and geographers are only scarcely utilized yet. Resulting shortcomings and biases in the context of bicycle safety research can be found on the level of concepts, data, analysis design and application. Conceptually many studies ignore the spatial dimensionality of transportation phenomena. Okabe et al. [9] and Okabe and Sugihara [10] extensively elaborate on the implications of disregarding the network-bound character of most transportation issues. Bicycle accident analyses, such as kernel density estimations (KDE) for the identification of hot-spots, must consider the dimensionality: in an urban environment accidents can only occur along roads; thus the models and algorithms need to be adequately adapted. There is also a spatial notion on the widely discussed data issue (see [11] for an overview). For many analyses statistical data are used which were aggregated on completely different levels [12], often at different points of time. An impressive example for this problem has been recently published on the author’s weblog [13]. Beside geospatial conceptual and data-related issues, geography plays a crucial role in the analysis design. Vandenbulcke [12] and Schepers [14], for example, note that the spatial autocorrelation (see [15] for further explanations) is not adequately considered in many bicycle accident analyses. The majority of standard statistical routines rely on independent samples; a requirement spatial phenomena, such as bicycle accidents, can hardly ever meet [16]. Although geospatial applications for “vélomobility”, as Koglin [17] summarizes all bicycle-related mobility issues, have been developed for quite a long time (see [18] or [19] for example), most of these efforts are isolated and very specific for a certain demand. There are only very few examples where a GIS is employed as integrated platform for more extensive application purposes. Larsen et al. [20] build a planning tool for bicycle ways upon geospatial analysis and Rybarczyk and Wu [21] combine geospatial analysis with multi-criteria decision analysis for building new bicycle facilities. Milakis and Athanasopoulos [22] use GIS not only for analysis purposes but also as geo-enabled communication platform for a participatory planning process. With regard to bicycle safety there is hardly any example where the results of a geospatial investigation of bicycle accident occurrences is fed into models which build the backbone for various consecutive analyses and applications.

The motivation of this paper is to showcase the benefit of such an integrated, geospatial perspective on bicycle safety. Geographical information systems (GIS) can in fact serve as facilitator for a holistic understanding of complex systems [23], such as bicycle safety, and form the backbone of interrelated analyses and applications. Based on two use cases, each with several
practical examples, the additional value of an explicit spatial perspective on bicycle safety should be exemplified and highlighted.

3 USE CASE I: SPATIAL ANALYSIS OF BICYCLE ACCIDENTS

Knowing where and when bicycle accidents happen is of crucial importance for targeted counter measures. As Nordback et al. [24] note, a more evidence-based approach to traffic safety is required in order to prevent none-appropriate measures. The analysis of past bicycle accidents helps to detect patterns, such as spatial hot-spots or temporal variations. Based on this, hypothesis about underlying processes and correlations can be formulated and tested, before counter measures – be it for example physical infrastructure measures or road surveillance – are taken.

![Figure 1](image)

Figure 1. From bicycle accident analysis to targeted counter measures for safer cycling. This workflow can be implemented in a GIS and applied for an initial analysis and for monitoring purposes.

For the following examples an extensive bicycle accident database is analysed. The dataset comprises of over 3,000 geo-located bicycle accidents in the city of Salzburg (Austria) between January 2002 and December 2011. The data come from police reports which are further processed by the national bureau of statistics and the responsible road maintainer. In the present case the data are provided by the department of urban and transport planning of the city of Salzburg. Although the data are of high, inherent quality in terms of spatial and temporal accuracy and degree of detail, the issue of severe underreporting (extensively treated in this OECs report [11]) has to be considered before final conclusions are drawn.

In a first step the spatial and temporal variation of bicycle accidents is investigated before the accident occurrences are related to simulated bicycle flows, allowing for a rough estimation of accident risk on the smallest possible scale of an urban road environment.

3.1 Spatial dynamics of bicycle accidents

Analysing the spatial dynamics of bicycle accidents over time helps to detect potentially causal, underlying effects. Plotting geocoded accidents on a map and subsequently identify locations with a higher accident prevalence is a valid, initial point. Such maps allow the observer to intuitively frame the accidents in the spatial and temporal dimension and reason about covariate factors, such as road design or traffic volume. Hot spot maps can thus be regarded as a spatially informed perspective on the data. Observations made on this basis can then guide further analysis steps, of which numerous, well established methods do exist [16, 25].
Figure 2. Bicycle accident occurrences in the city of Salzburg (Austria). The spatial and temporal dynamics becomes obvious.

In the presented case (see Figure 2), 3,048 accidents occurred at 1,865 distinct locations. Within the ten years of observation 1,379 single accident locations are reported. At 16 locations ten or more accidents are reported (Figure 3). These hot-spots account for 199 (6.5%) accidents and are of special interest for further investigations. All of these hot-spots are located at intersections, mainly at major radial roads which connect peripheral quarters with the city centre and the major bicycle connection along the Salzach river. Except hot-spot number 9 the multiple-accident locations occur on roads with unrestricted motorized traffic. Consequently the share of car-bicycle crashes is significantly higher (77% compared to 55%), compared to the overall statistic.

Figure 3. Spatial (left) and temporal (right) distribution of accident hot-spots in the city of Salzburg (Austria).

The right chart in Figure 3 clearly indicates that the temporal variation of accident occurrences at the hot-spots is relatively homogeneous. Thus it can be concluded that the accident occurrences at these locations are not due to temporal obstacles, such as construction sites, but rather due to an inappropriate road design.
The exploration of all bicycle accidents on different spatial and temporal scale and aggregation levels reveals relevant variations. As it is demonstrated in Figure 4, the total number of accidents is spatially clustered in the city centre along the Salzach river and major radial feeder roads. It is important to note, that these figures don’t allow for any deduction of risk factors. Still, they help to identify areas with a comparable high number of accidents, which are in turn areas where targeted counter measures can be expected to have a high impact.

**Figure 4.** Spatial pattern of accident occurrences: absolute number of accidents per grid cell (left) and accidents per road kilometre.

**Figure 5.** Temporal dynamics of accident occurrences. Charts on the right show aggregates of ten years with (from top to down) month, weekday and hour as aggregation level.
In the temporal dimension seasonal variations and reoccurring patterns in the course of a week and a day can be clearly observed (Figure 5).

Despite the imperfect data availability (underreporting, non-representative spatial distribution of counting stations), the data suggest a high degree of dependency between the number of accidents (aggregated for the ten years of observation) and the bicycle traffic volume (counting data for one year at one central location). For weekdays a correlation coefficient $r = 0.984$ and for hours of the day $r = 0.977$ are calculated. The correlation between bicycle traffic volume per month and accident occurrences is $r = 0.865$; this can be, at least partly, explained by the location bias of the used counter. Nevertheless, for a sound estimation of the risk on the level of road segments, bicycle flows for each segment are required as input data.

### 3.2 Estimating bicycle accident risk

Risk analyses for bicycling generally follow one of these two approaches: epidemiological analyses use aggregated accident data which are related to exposure variable, derived from statistical data, such as inhabitants or distance travelled per reference unit [26, 27]. Reference units are typically scaled between census districts and national regions. In contrast to this macro-scale approach, in-depth analyses of individual accidents aim to reveal underlying risk factors which are then used to model the risk on a higher scale level [5, 28]. Both approaches help to better understand the spatial dimension of bicycle safety. Nevertheless the quality of the required data must be of high quality to generate sound results, what is in fact a critical flaw in many studies as extensively noted by the ITF-OECD Working Group on Cycling Safety [11].

Currently there are several methods to measure or estimate the volume of bicycle traffic. Lindsey et al. [29] give an overview of common technical hardware, such as induction loops and infrared sensors, for bicycle and pedestrian counting at fixed locations. Other studies estimate the number of bicycles based on travel to work data from censuses [30], which of course can only be a rough approximation. A rather new approach is to take crowd-sourced data which is sensed by mobile devices and published in social media applications. Griffin and Jiao [31], for example, use data from the fitness app Strava to estimate bicycle flows. What all these three approaches have in common is that the data are neither complete on the scale level of single road segments nor representative for the whole population. It is thus a significant challenge to model bicycle flows appropriately on the road level in order to have sound exposure data available for risk calculations.

Different to the three approaches mentioned above, the combination of geographical information systems with agent-based models allows for a global estimation of bicycle flows on the most detailed scale level. As Wallentin and Loidl [32] demonstrate, average flows for each segment in a road network can be accurately modelled based on individual behaviours of multiple different agents. In the following risk estimation the data from Wallentin and Loidl [32] are used as exposure variable for an estimation of risk on the level of road segments.
Figure 6. Accident risk estimated from simulated bicycle traffic volumes and reported accidents. The estimation can be done for different reference units: spatially aggregated to regular hexgrid cells (left) or related to individual road segments (right).

Compared to Figure 4 the patterns emerging in Figure 6 are significantly different. Whereas the absolute number of accidents is highest on the main bicycle connection along the Salzach river, the risk of getting involved in an accident is comparable low. The clusters of high risk that emerge in the Southeast and North of the city are most probably due to high volumes of motorized traffic in these areas. A closer analysis of the single road segments (see right map in Figure 6) reveals an important relation: the lower the quality of the bicycle infrastructure is (in the worst case roads without any bicycle infrastructure) the higher the risk for accidents gets. Although this risk calculation is based on models (which are by definition an abstraction of reality), the analysis in the present case study is another [33, 34] evidence against the “vehicular cycling” paradigm [35, 36].

Once the correlation between the built environment and accident risk is proven, it can be employed in global assessment models. Such models allow for the assessment of the road quality in terms of safety and can be utilized in various applications, ranging from planning tools to information systems.

4 USE CASE II: MODELLING SAFETY THREATS

A geographical information system is a suitable environment to bring together various perspectives on the road space and extract additional information from this overlay. In the context of bicycle safety GIS can thus be used to relate multiple descriptions and objects of the road space to each other and use these variables as inputs for complex models which facilitate integrated analysis approaches [8].

Such a model is proposed by Loidl and Zagel [37]. It compiles results from accident analyses (see previous section), findings from literature, expert interviews and user feedbacks in an indicator-based assessment model. The basic idea behind this approach is the following: variables (“indicators”) which contribute to the risk for bicyclists are identified, weighted and compiled in a spatial model. This model is than applied on a whole road network, digitally represented in a GIS. The indicator-based assessment model results in a standardized index.
value which indicates the quality of a road segment in terms of road safety (ranging from 0 = excellent, no risk to 1 = bad, high risk). The relation between the quality of the road space and the accident risk is shown in Figure 7.

![Figure 7](image.png)

**Figure 7.** Scatterplot of accident risk and index value, calculated by the indicator-based assessment model. The map (right) shows the spatial distribution of the index value (average per hexgrid cell).

The assessment model is completely adaptable and thus open to changes of the indicator set and the respective weights. The default composition is based on an iterative calibration process (for details see [37]) and applied in the context of an operational web application (www.radlkarte.info). In the following the applicability of the indicator-based assessment model and, connected to that, the advantage of spatial modelling in a GIS is demonstrated in three different application examples: status-quo analysis and weak-point detection, simulation of measures and routing.

### 4.1 Status-quo analysis and weak point detection

Using the indicator-based assessment model [37] for the purpose of status-quo analysis and weak point detection has at least two advantages compared to other evaluation or assessment routines:

a) The resulting index value can be traced back to the input data. This ensures a maximum transparency and facilitates a better understanding of number values, independently from the evaluator.

b) The standardized index value allows for a sound comparison of different areas within or between regions under investigation.

The analysis can be done on the level of single road segments or for aggregated reference units (as it has been done for the map in Figure 7). Additionally the assessment can be extended to pre-defined routes, indicating the quality of accessibility of facilities.
In Figure 8 the catchment area of the faculty of natural sciences (University of Salzburg) is investigated in terms of bicycle suitability (“bikeability”). The quality of the road network was assessed within 1.5 kilometres and aggregated to hexgrid cells (left part of Figure 8). Based on this analysis a further abstraction step reveals the most significant problem in this area: although the immediate surrounding of the university building is equipped with excellent bicycle infrastructure (mostly separated bicycle ways) and the high-quality bicycleway along the Salzach river is not far away, the two corridors are not connected adequately.

With the indicator-based assessment model such analysis can be done on various scale levels. They are relevant inputs for evidence-based planning processes and help central facilities, such as university buildings, schools, hospitals etc., to enhance their accessibility for active modes of transport. As Garrard et al. [38] and Parkin et al. [39], among others, have demonstrated, the accessibility and proximity of safe bicycle infrastructure is of crucial importance for the acceptance of the bicycle as a safe, utilitarian mode of transport.
The accessibility of central facilities does not only depend on the immediate surrounding. The route from an origin to the destination might be without significant obstacles but still not attractive in terms of safety. In order to assess the accessibility of the faculty of natural sciences (University of Salzburg) routes from 1,500 origins are calculated. The index value of the traversed segments is then averaged and mapped in an origin-destination map (Figure 9). As it can be seen, the accessibility of the university building is relatively good from origins in the Northwest and Southeast. A lack of adequate connections can be found for origins in the Northeast.

The given examples are commonly applied on the macro- and mesoscale level. In the next section, the indicator-based assessment tool will be applied on the road level in order to detect local weak points and simulate various measures for a safety improvement.

4.2 Simulation of measures

Geographical information systems are not only suitable for the digital representation of the status-quo at a given point of time but facilitate dynamic interactions [7]. Thus they can be used for scenario development and simulation purposes. As stated before, the indicator-based assessment model is designed in a way that it allows for a dynamic adaption of indicators and the respective weights. Hence it can be applied in a dynamic environment, where, for example, the effect of physical changes in the road space on the overall quality is simulated interactively. For this purpose a web tool, based on Loidl and Zagel [37], was developed by Wendel [40].

![Figure 10](image.png)

Figure 10. From weak point detection to the simulation of possible interventions.

Figure 10 demonstrates the benefit of this interactive simulation environment. Disruptions in continuous corridors of high quality bicycle infrastructure become obvious in the visual interface (a). The associated attributes which describe the road space at the respective segment explain the index value (b). In the presented case the two road segments with a high index value, indicating a comparably high safety threat, lack of any bicycle infrastructure. In a third step the effect of several measures, such as building different types of bicycle infrastructure or restricting motorized traffic, can be simulated (c).

The simulation of measures in a digital environment is of high value when it comes to cost-effect optimization and budget allocation. The immediate visual feedback and update of the index value help decision makers to simulate various options very fast and finally come up with evidence-based decisions which can be communicated transparently.

4.3 Routing

As demonstrated in the previous two application examples the spatial modelling of safety threats for bicycles can be an informative building block of integrated planning processes. Additionally the model results can also be employed in information systems for bicyclists. Road
safety is a key value for existing and prospective bicyclists [2, 41]. Beside the provision of safe infrastructure information about optimal routes is a major part of integrated bicycle promotion strategies [42, 43]. Most routing information systems are designed for automotive traffic. There the shortest or fastest route equals the optimal route in most cases. Different to this, route optimization for bicyclists is more complex as various decision factors, above all safety concerns, need to be addressed.

Figure 11. Route optimization based on a spatial assessment model. The connection between the index value from the assessment model and the route optimization becomes obvious (right). The same optimization is applied in the public bicycle routing application www.radlkarte.info, where the optimal is compared to the shortest route (left).

Using the index value resulting from the indicator-based assessment model as impedance in a routing engine calculates optimal routes for bicyclists. The optimal ("safest") route is a trade-off between the index value and the distance, in order to calculate reasonable route recommendations.

5 CONCLUSIONS

Transportation systems and mobility, and thus bicycle safety, are spatial phenomena by their very nature. Hence the employment of geographical information systems (GIS) as integrated platform for multiple application purposes helps to gain relevant insights and extract useful, additional information.

The geospatial perspective on bicycle accidents reveals distinct patterns and dynamics in the spatial, but also in the temporal dimension. Together with advanced model approaches, such as agent based simulation, the geospatial accident analysis even allows for the estimation of risk factors on the most detailed level of single road segments. Gained insights from spatial analysis of bicycle accidents, combined with additional information sources can then be compiled in geospatial models, such as the indicator-based assessment model. The flexibility and global applicability of this model makes it a useful element of innovative analysis and planning tools as well as an algorithmic backbone of routing optimization.

In the best case, all these efforts contribute to comprehensive strategies for safer cycling. Starting from status quo analysis, planning processes and counter measures can explicitly address the safety issues through geographical information systems. Additional to the best possible infrastructure, that is safe roads, spatial information systems, such as routing applications, actively contribute to accident prevention and an increased attractiveness of the bicycle as safe mode of transport.
REFERENCES


