Quantitative, Qualitative, and Historical Urban Data Visualization Tools for Professionals and Stakeholders

Cody Dunne¹, Carl Skelton^{2(X)}, Sara Diamond², Isabel Meirelles², and Mauro Martino¹

¹ IBM - Watson Health - Cognitive Visualization Lab, Cambridge, USA {cdunne, mmartino}@us.ibm.com
² OCAD University - Visual Analytics Lab, Toronto, Canada carl@ultratopia.com, sdiamond@ocadu.ca, imeirelles@facultv.ocadu.ca

Abstract. Existing technologies for transportation planning, urban design, and decision-making have not kept pace with rapid urbanization. Visualization and analysis tools can help by combining qualitative, quantitative, and historical urban data – helping experts understand the *system of systems* of the modern city. Incorporating insights from experts in several relevant fields, we have derived a performance specification for visualization tools supporting general transportation planning problems. We examine two existing technologies against the specification – Betaville and StoryFacets – and recommend adapting them as first-generation urban system analysis/planning support tools. We also suggest guidelines for the next generation of tools for transportation planning.

 $\label{lem:words: Visualization · Quantitative data · Qualitative data · Historical data · Transportation planning · Urban systems · Information technology$

1 Introduction

Over half of the world's population (more than three-quarters in developed countries) now lives in urban areas, and people are predicted to continue migrating to cities over the next several decades. Systems analysts, policy makers, designers, and citizens require appropriate visualization and decision-support technologies for planning, design, and decision-making. Thanks to the ubiquity of networked mobile commercial and personal devices (and the skills to use them in new ways), we now have at our disposal massive, dynamic data sets. Using cloud-based data aggregation and processing power, we can merge this in real time with historical data like maps, land use, and demographics.

Connecting these data sets is valuable, but it is only part of the solution. For planning and design purposes, "hard" data is much more useful when properly associated and correlated with qualitative information–attitudes and preferences and priorities, as well as the dynamics of inhabiting and circulating within urban environments. Existing urban system visualization, operations, planning, and design tools do not yet afford such integration.

In this paper, we posit that it is now possible to bring together quantitative (engineering) data and qualitative (design/experiential) information in a full-spectrum knowledge space. As an initial instance of this approach, we examine transportation

© Springer International Publishing Switzerland 2016 N. Streitz and P. Markopoulos (Eds.): DAPI 2016, LNCS 9749, pp. 405–416, 2016.

DOI: 10.1007/978-3-319-39862-4_37

infrastructure, a particular class of urban systems for which quantitative and qualitative information sources are abundant, and often in controversy. Our discussions with domain experts have revealed two key research problems for which visual analytics would be particular beneficial: evaluating *complete streets* and *parking management*. We present a set of performance specifications derived from these case studies which we believe are generalizable to public information works beyond transportation.

We believe fast progress can be made on transportation problems through the adaptation and innovative combination of readily available tools, embodying different approaches to quantitative/qualitative representation – the **Betaville** massively participatory online platform and the **StoryFacets** visual data exploration system. We present a comparative evaluation of Betaville and StoryFacets as first-generation urban system analysis/planning support tools, and make recommendations for future research extending and exploiting them to meet the performance specifications.

2 The Big Picture: Public Information Works

Open data, in and of itself, is not the same thing as *open information*. An *open information* policy would imply the further duty to provide data in citizen-accessible and citizen-intelligible form, and to provide for public access to the information used by experts, decision-makers, and service providers acting on urban systems.

We propose a new approach to urban IT support systems and data resources – that in the aggregate, they be re-conceived as public works in the traditional sense of that term – large structures built and maintained for general public use, i.e. both by professionals whose work impinges on the built public realm and as public (citizen) information resources. Within this conceptual framework, an urban IT infrastructure is implemented less like a boiler or sewer system (in the dark, only to be seen and handled by staff), and more as a medium of exchange between staff, policy-makers, proponents, designers, and citizens – within the conceptual framework of systematic cultivation and support of new levels of citizen expertise, as well as engagement.

Public IT/information spaces must provide for a well-articulated set of discrete user interaction profiles, from novice to expert, including provision for "leveling up"— i.e. self-directed education and skill development, from basic web literacy up to levels of expertise currently only accessible to professional specialists. At maturity, such a system would amount to a common *back end*, with enough discrete purpose- and user-type-specific *front ends* to fully gather, process, understand, and communicate the urban systems we have, and to make sense of how they can and should co-evolve in the future. Our current work is to take viable first steps toward this goal in the domain of urban transportation infrastructure.

3 The First Application Space: Transportation

In strict engineering terms, a transportation network can be considered as one of the systems in an urban "system of systems" along with emergency response, energy, water,

waste disposal, health, law enforcement, education, and so on. For some of the subsystems in the transportation network, like highways and subways, citizens and engineers are traditionally concerned with similar (quantitative) issues of system performance: maximizing throughput, speed, and consistency. For others, like shopping and residential streets, quantitative measures of transportation system performance do not adequately characterize or address overall urban system *quality*, as experienced by real people in what they experience as urban environments. Concomitant with the goal of converting open data to public information, we propose upgrading urban system informatics from strictly quantitative data to the full spectrum of quantitative and qualitative information, to address the operation and improvement of transportation networks as "lived systems", i.e. environments.

This concept is being explored in the context of "iCity: Urban Informatics for Sustainable Metropolitan Growth" a translational public-private partnership research initiative led by the University of Toronto Transportation Research Institute (UTTRI). iCity includes participants from a variety of other research units at University of Toronto, OCAD University, University of Waterloo, as well as public agencies and industry partners including the planning departments of the cities of Toronto and Waterloo, and industry partners IBM, ESRI and Cellint.

By leveraging insights from experts in computer science, transportation engineering, urban planning, visualization, and real-time data analysis we hope to define strategies for the development of next-generation tools/infrastructure for better-quality – rather than simply more efficient – transportation operations, planning, and design.

Discussions with experts in transportation planning have yielded two research problems for which next-generation visual analytics could be particularly beneficial: the relatively technical domain of *parking management*, and the more complex set of qualitative/quantitative factors implicated in the *complete streets* policy guidelines. Below we detail the background and goals of this research, which we leverage in a preliminary performance specification for integrating visualization tools into public information works.

3.1 Use Case: Parking Management

Urban parking has a significant impact on traffic congestion and behavior [3, 6]. While on- and off-street parking policies, alternative pricing models and smart parking technology can play a substantial role in reducing urban congestion, cities have been slow to adapt. Existing research has addressed the parking problem from two perspectives: (a) analysis of the relationships between parking supply, demand, and the incidence of illegal commercial vehicle parking [12]; and (b) development of a traffic simulation tool that incorporates driver decisions of parking space choice, and simulates the effects of parking search patterns on traffic congestion [4].

Our collaborators are working to expand upon these efforts to develop holistic, fully functional, operational tools for the management of parking in congested urban areas. Of particular interest is the context of supply and demand of on-street parking, road congestion, transportation networks and traffic flows, parking by-laws, smart parking, and pricing models. Methodologically, they seek three different orders of data to support

their work: observed parking behavior and effects on the existing grid, with its current regulations; simulated behavior and impacts using agent-based model simulations, typically represented as schematic 2D animations of rectangular cars; and variable-parameter first-person immersive 3D parking "games", in which real commuters navigate and make choices according to actual and hypothetical scenarios.

3.2 Use Case: Complete Streets

Complete Streets [9] is a design/policy framework for the configuration of city streets to provide for a full spectrum of users, with an unusually rich definition of system performance: "complete" streets are "designed for all ages, abilities and modes of travel. Safe and comfortable access for pedestrians, bicycles, transit users and the mobility impaired is not an afterthought, but is an integral planning feature" [4]. Our collaborators are working on tools to rationalize the conversation about complete streets in the context of information technology, and to make explicit the assumptions and trade-offs that are implicit in street design. In particular, they wish to:

- Quantify the benefits and costs of alternative street designs, including: emissions
 exposure, travel delay, access to facilities, physical activity, and conflicts between
 pedestrians/cyclists/parked vehicles/transit vehicles.
- Model complete streets with inputs for mode and purpose demands, existing constraints such as right-of-way and built environment, technology for space sharing, and user-defined priorities. The model will recommend optimal right-of-way space use to balance competing needs (number of vehicle, dedicated transit, and bicycle lanes; pedestrian walkway width; social space; parking; cyber systems).
- Integrate complete streets models with travel demand models for auto, commercial vehicle, transit, active transportation, and parking needs.

3.3 Initial Specification

From these two use cases we've identified a preliminary performance specification for visual analytics tools to support transportation planning tasks, as sample "Public Information Works" components:

Design approach

- Communication-centered. Analysis and visualization tools should be integrated such that consumable visualizations for both specialist and non-specialist stakeholders (citizens, leaders, proponents) with varied needs can be generated.
- Collaboration-minded. Tools should be designed to support teams and groups interacting on analysis projects synchronously (real-time chat, shared workspace) and asynchronously (comments, versioning).

Visual data/model integration

• Qualitative data. Systems should allow association of statistical with experiential information. Tools should include stakeholder feedback through channels such as social media, and link documentary media (text, photos, audio, video files) to interactive map-based or immersive (3D fly-through) infographics.

- Real-time "what if" scenario support. Specialist users require integrating (and explicitly displaying) mathematical models with powerful analysis tools and statistics for simulation/scenario development on-the-fly. Users must be able to interactively change model inputs, including user-defined priorities, and visually understand the effect on model outputs.
- Changing/historical data and data ontologies. Available data will change over time (surveys, utilization monitoring, and cell phone paths) and its evolution should be recorded and visualized. This is valuable for understanding both snapshots of history and overall temporal behavior.
- **Provenance.** As data is processed by user filtering actions and modeling, it is valuable present the history of workflow visually to users, so as to help them create more repeatable, faster, and accurate analyses.

Visualization techniques

- **Interactive computing.** Tools should support interactive data exploration, including user manipulation (select, filter, zoom, join, model parameter changes).
- Overview + detail. Visualization tools should provide an overview of the data or model (e.g., all of Toronto), as well as tools to drill down into local detail.
- **Geospatial visualization.** Map-based views of scenarios and statistics should be included for understanding inherently geospatial transportation systems.
- Information visualization. We recommend using spatial layout to encode attribute values to complement geospatial visualizations. Patterns in model outputs such as price, availability, congestion, and exposure can be easily understood when using spatial encodings.
- Comparative visualization. Tools should support visually comparing model results from different inputs, outcomes of proposed scenarios, and changing or historical data snapshots to enhance understanding.

- White boxes

- Ontology. Any first-generation system must clearly display its ontology as a set of hypotheses. This includes explicit and accessible definitions of entities, properties, and relationships according to which any given sensor input is construed as information, recorded, and processed by software systems, to address both quantitative and qualitative aspects of data characterization and processing.
- Models. The mathematical models of system behavior that underlie system simulations must be accessible from within the user environment, as matters of public record/understanding.
- **Provenance.** Tools should maintain retrievability of each version of the data and the workflow which created it.

Over the course of the next four years, through an iterative cycle of development and evaluation, we will build and deploy applications to support transportation system operations, optimization, and planning as useful products in the near term, and proofs-of-concept for the broader "Public Information Works" approach down the road.

4 Technology in Hand

Two current complementary technologies together embody many of the aforementioned performance specifications for our case studies, and can be further enhanced to meet transportation planning needs: the Betaville massively participatory online platform and the StoryFacets visual data exploration system. We believe that fast progress in urban system development can be made by using these tools as a starting point.



Fig. 1. The Betaville desktop client, "God's Eye" view

4.1 Betaville

Betaville (Fig. 1) is an online environment for distributed development and deliberation about possible changes to built environments, from the scale of a public artwork to that of district-level urban development or re-development. It is intended to provide for adhoc online exchange of ideas between stakeholders, proponents, and professional experts, with a view to engaging all three user types proactively in "pre-design": identification of key issues, and informal putting forward of sketch/schematic models for discussion and elaboration into rough-but-robust concepts in advance of formal design development and approval processes.

Design approach – Communication-centered and Collaboration-minded: The idea for Betaville was originally developed to provide for timely and constructive modes of public engagement [8]: rather than waiting for public or private-sector proponents to present proposals *after* large investments in detailed technical design and process overhead had already been made, stakeholders could engage in ideation and problem-solving

directly: not just commenting, but also putting forward their own ideas for all or part of a project, for further deliberation and elaboration in a persistent shared online design space. At a minimum, proponents could have ready access to current information about community interests before proceeding with conventional proposal development; in the best cases, new coalitions could form online, innovating more freely in the early stages of project ideation, building new constituencies and partnerships, building consensus as viable concepts come to maturity.

Visual data/model integration – *Qualitative data:* Betaville's qualitative value lies primarily in its ability to represent simplified models of proposed projects in a recognizable context, either through augmented reality or an immersive 3D model in a desktop "game" application, linked to metadata and external web resources (Fig. 2).



Fig. 2. The Betaville "Citizen's Eye" view of the same scene as Fig. 1

Visualization techniques – Interactive computing: Additional information can be represented in pop-up windows by clicking on objects in the scene, external web links, or infographic overlays. For example, showing the predicted change in emissions for a given scenario.

- Overview + detail: The key experiential aspect of Betaville is that it provides for the kind of God's Eye View generally associated with urban planners and chambers of commerce, and the immersive "first-person" citizen's perspective.
- Geospatial visualization: As a 3D visualization of real and proposed worlds, the main view of Betaville is inherently geospatial.
- Information visualization: Elements in the geospatial visualization are linked with quantitative data, which can be visualized on demand.

White box – Ontology: The Betaville environment is self-documenting in real time: all proposal versions and comments are user-signed and time-stamped, so concept development and discussion history can be retrieved for review and analysis.

4.2 StoryFacets

StoryFacets [5] is a visual exploration system for relational data, which is particularly suited data analysis such as transportation planning where collaboration and communication with stakeholders is key. It is shown with alterations in Fig. 3.

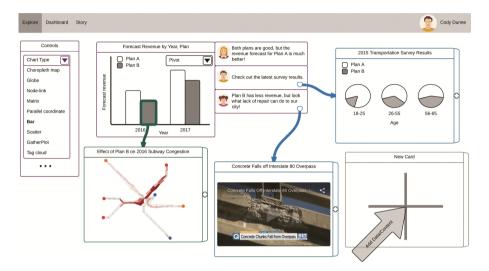


Fig. 3. Mockup qualitative extension to the StoryFacets trail view incorporating visualizations alongside annotations and comments with text labels, images, videos, and webpage mashups.

Design approach – Collaboration-minded: StoryFacets was built as a web-based visual analytics system, to reduce barriers to entry and facilitate collaborative analysis and sharing. Moreover, it includes user- and project-management features for distributed teams and citizen scientists to collaborate effectively.

- Communication-centered: Eventually insights must be communicated to stakeholders – such as colleagues, managers, customers, or even the general public – to be useful and actionable [10]. However, many analysis tools still suffer from the so-called "Power-Point gap:" analysts often end up copying and pasting screenshots of tables and tools into a PowerPoint slideshow. Moreover, appropriate design for communication depends on analyst expertise, the presentation's audience, and the nature the presentation. E.g., an exploratory visualization system designed to support complex transportation model analysis is not necessarily the best way to present policy recommendations to the general public.

To ease communicating visual insights from analyses we designed StoryFacets as a one-source, multiple-media exploratory visualization presentation tool. Here, the

underlying data and analytic provenance model is shared across many linked visualizations. The trail view presents visualizations and their provenance, the dashboard view supports high-level analysis of visualizations and other content in a space filling layout, and the slideshow view enables step-by-step storytelling. Changes in one view are instantly reflected in the others, eliminating the error-prone conversion between analysis and presentation.

Visual data/model integration – Provenance: Exploring and understanding complex relational data often require several sessions, and when returning to a previous analysis it can be difficult for users to recall the steps in their workflow. A data scientist iteratively cycling through tasks can easily forget the exact steps already done and mistakenly omit or inconsistently perform operations. For example, a transportation planner could use a filter to analyze only rush hour traffic on weekdays, create a visualization showing street throughput, and then discover a neighborhood missing data. After performing a new data collection and integration, they could accidentally omit the weekday filter and create a new visualization with a different overall message.

Exposing users to their analytic and data provenance enhances recall between sessions [2, 7, 11], as well as analysis comprehension [1]. However, most visualizations present users with only an undo stack. A handful of tools present separate hierarchical history views, and in general require extra maintenance effort from users.

StoryFacets exposes provenance embedded directly within the analysis workspace to enable easy understanding and sharing of results. During exploration, user interactions leave a trail of visual and textual bread-crumbs which document the reasoning process and data provenance. Exposing the analysis process visually enables users to utilize the spatial memory and track specific interactions. Using this design led to increased insight discovery, users could recall their findings and the exact process used to arrive at them, as well as interpret explorations of others [1].

- Qualitative data: In StoryFacets, visualizations can be augmented with cards that allow annotations and qualitative content to be added, such as text captions, bullet points, hyperlinks, images, video, and even interactive webpages. These cards use the Markdown markup language which is easy for new users to pick up. Users are given a text editor to enter Markdown with a real-time preview alongside. Upon submission, the card displays the rendered Markdown and includes an edit button to display the text editor again. For example, a user may add a video of a bridge collapsing due to poor maintenance to an analysis of infrastructure spending proposals.

Visualization techniques – Interactive computing: Common analysis operations like filtering are done through simple drag-and-drop interactions which create new cards. Simple user interface widgets are used to provide a pivot mechanism to transition between linked aggregates, which allow fluid exploration of multiple node or edge types. These interactive exploration mechanisms are quickly picked up by novice users [1], and are supported by quick system response times.

- Overview + detail: In StoryFacets, visualizations are displayed on a zoomable and pannable canvas, connected by parent links which expose the data provenance and

exploration history. This interface allows users to get an overview of the entire exploration, as well as zoom into analysis of subsets of the data. Moreover, the filter and pivot mechanisms allow users to begin with an overview visualization which they trim into a meaningful subset to answer their question.

– *Information visualization:* StoryFacets focuses on aggregating visualizations which provide visual scalability for large data sets. The modular architecture allows easy integration of additional visualization types, but we include a general-purpose visualization called a GatherPlot. This extension of a traditional scatterplot stacks items with the same value together for easy countability and distribution analysis.

White boxes – Ontology and Provenance: StoryFacets offers not only a cross-referenced catalog of all relevant information assets, it also tracks and displays individual and aggregate user itineraries through those resources, a meta-mapping of how users navigate in the information space, providing for new levels of analysis of which resources are sought out, how they are discovered and used, how they are associated with each other in use... in other words, elements of an "ontology in use", which should inform subsequent tool development, data collection, and in due course development of the built environment itself.

5 Discussion and Opportunities for Future Research

Our ultimate purpose is to define strategies and guidelines for the development of the next generation of tools/infrastructure for transportation operations, planning, and design – the transportation-specific components of a comprehensive urban IT infrastructure for cities worth serving: efficient, sustainable, adaptable, and desirable. As the city itself is a human system of human systems, so its information works must develop as a support system of support systems.

Betaville and StoryFacets approach the association of very different qualitative/ quantitative data-to-information assets, from within two very different application/interaction genres. Combining those two perspectives within a smart model, i.e. a world in which individual objects can be linked to city data or external web resources, and providing for iterative development of design proposals and forums about them, with ID and time-stamping of individual contributions, provides in principle for something like a permanent open "charrette", exactly the kind of logistics and protocol that have supported the development of open-source software generally.

Betaville and StoryFacets will be adapted to the purpose of proving this to be a practical proposition in the long term, and bringing it to the service of transportation operations and planning, within the iCity consortium. For the applications of Betaville and StoryFacets to complete streets and parking management, there are several missing components of our performance specification which pose open research problems:

Design mindset – Communication-centered and collaboration-minded: In theory, the complementarity of Betaville and StoryFacets would suggest another mash-up, to combine their disparate functions and virtues in a single unified environment. The

working assumption would have to be that an adequate critical mass subset of stakeholders within any given community will have the tools and skills (personal computer, internet access) to operate the client, and sufficient aggregate motivation to contribute either to general community awareness and understanding of active development projects, or ideally to help build an ongoing local culture of ideation and deliberation about issues, possibilities, potential improvements over the long term.

Visual data/model integration – Qualitative data: We propose to build out StoryFacets as a browser for disparate asset types (Fig. 3), from public GIS data to real-time traffic mapping, social media, survey data, graphics, photos, audio, and video – incorporating these into our ontology.

- Real-time "what if" scenario support: As more models are developed for transportation planning, it will be necessary to integrate both interactive controls for and the results of these algorithms. Careful design will be needed to extend our existing tools to incorporate these models, yet retain usability.
- Provenance and changing/historical data: As the underlying data changes, systems must be able to record these changes and present them to the user visually, in addition to allowing the user to switch between "snapshots" of the data. The StoryFacets model is well suited to such an extension.

Visualization techniques – Comparative visualization: Particularly in conjunction with incorporating model results and changing data, it is critical to visually show comparisons. E.g., when adjusting model parameters for bikeshare use, a direct comparison of emissions across multiple scenarios is more accurate than viewing them separately. We will integrate both interactive and directly comparative techniques.

White boxes – Models: In addition to "what if" scenario support, in which model inputs and outputs are integrated with tools using interactive computing techniques, the algorithms behind those models must also be exposed as much as possible.

- Ontology: As a data structure visualization tool, StoryFacets is readily adaptable to the purpose of representing (visualizing) the computational ontology underlying the data/information being navigated and processed, even as that ontology evolves to account for new data resources, new policies, and new methods in transportation network operations and planning.

6 Conclusion

The goal of integration of qualitative with quantitative dimensions of transportation system planning is already ambitious; the deeper integration it calls into question, of integrating research and professional service-delivery infrastructure with stakeholder communication and engagement platforms, is even more so. In this context, the conventional distinction between user-friendliness and expert-friendliness breaks down, as professionals and lay stakeholders converge: a need for visualization, analysis, and

simulation tools that support continuous "up-skilling" in understanding and exploitation of scales and orders of data and information that are themselves evolving, in real time. As we build out and adapt Betaville and StoryFacets as complementary proofs-of-concept for this approach in transportation applications, we expect our research and development work to be exploited not only for transportation applications, but in due course to integrate the full range of urban system analysis, operations and planning in the service of cities that are not only smarter, but wiser, through the active cultivation and support of more (and better-informed) engagement by any combination of citizens, leaders, and professionals working together, 24/7.

Acknowledgments. This work is supported by the Ontario Research Fund – Research Excellence Round 7, under the project iCity: Urban Informatics for Sustainable Metropolitan Growth and MITACS.

References

- Dunne, C., Riche, N.H., Lee, B., Metoyer, R.A., Robertson, G.G.: GraphTrail: analyzing large multivariate, heterogeneous networks while supporting exploration history. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI 2012, pp. 1663– 1672. ACM, New York, NY, USA (2012)
- Lipford, H.R., Stuke, F., Dou, W., Hawkins, M.E., Chang, R.: Helping users recall their reasoning process. In: Proceedings of the IEEE Symposium on Visual Analytics Science and Technology, VAST 2010, pp. 187–194 (2010)
- Miller, E.: First we take Toronto ... lessons from the TASHA implementation and next steps.
 In: Behavioural Detail and Computational Demands in Agent-Based Models Workshop (2015). http://www.fcl.ethz.ch/assets/EJMiller-ABM-Workshop-2015.pdf
- Nourinejad, M., Wenneman, A., Habib, K.N., Roorda, M.J.: Truck parking in urban areas: application of choice modelling within traffic microsimulation. Transp. Res. Part A Policy Pract. 64, 54–64 (2014)
- 5. Park, D.G., Dunne, C., Ragan, E., Emlqvist, N.: StoryFacets: generating multiple representations of exploratory data analysis for communication (2016, under submission)
- 6. Shoup, D.C.: The high cost of free parking. J. Plann. Educ. Res. 17(1), 3–20 (1997)
- Shrinivasan, Y.B., Van Wijk, J.J.: Supporting the analytical reasoning process in information visualization. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2008, pp. 1237–1246. ACM, New York, NY, USA (2008)
- 8. Skelton, C.: Soft City Culture and Technology: The Betaville Project. Springer, New York (2014)
- Toronto Centre for Active Transportation: Complete Streets for Canada (2012). http:// completestreetsforcanada.ca. Accessed 1 July 2014
- Viégas, F.B., Wattenberg, M.: Communication-minded visualization: a call to action. IBM Syst. J. 45(4), 801–812 (2006)
- 11. Ware, C., Gilman, A.T., Bobrow, R.J.: Visual thinking with an interactive diagram. In: Stapleton, G., Howse, J., Lee, J. (eds.) Diagrams 2008. LNCS (LNAI), vol. 5223, pp. 118–126. Springer, Heidelberg (2008)
- 12. Wenneman, A., Roorda, M.J., Nurul Habib, K.: Disaggregate analysis of relationships between commercial vehicle parking citations, parking supply, and parking demand. Transp. Res. Rec. J. Transp. Res. Board **2478**, 28–34 (2015)