Between plan and reality: tracing the development dynamics of the Lanzhou New Area - a computational approach

Between plan and reality: tracing the development dynamics of the Lanzhou New Area - a computational approach

Master thesis for Advanced Urbanism
June, 2019

Author
Serjoscha Düring
117570

Examiners
Prof. Dr. Reinhard König
Prof. Dr. Sven Schneider

Institute for European Urban Studies
Faculty of Architecture and Urbanism
Bauhaus-University Weimar, Germany

Abstract: Contemporary planning practice is often criticized as too design-driven with a lack of both quantitative evaluation criteria and employment of models that anticipate the self-organizational forces shaping cities, resulting in significant gaps between plan and reality.

This study aims to introduce a modular toolbox prototype for spatial-analysis in data-poor environments. It is proposed to integrate designing, evaluation, and monitoring of urban development into one framework, thus supporting data-driven, on-demand urban design, and planning processes.

The proposed framework's value will exemplarily be tested, focussing on the analysis and simulation of spatiotemporal growth trajectories taking the Lanzhou New Area as a case-study - a large scale new town project that struggles to attract residents and businesses.

Conducted analysis suggests that more attention should be given to spatiotemporal development paths to ensure that cities work more efficiently throughout any stage of development. Finally, early hints on general design strategies to achieve this goal are discussed with the assistance of the proposed toolbox.

Table of Content

1. Introduction

Small settlements evolved step by step through villages to towns and eventually into cities. While most villages remained villages, towns primarily emerged at advantageousness places such as crossroads, fords, etc. - locations prone to serve as hubs for trade and exchange. Besides these organically grown towns, there are planned ones. the alluring idea of designing a whole city from scratch has been around for centuries. From the first known attempts in the Greek city of Miletus (six century BC), over the garden-city movement rooting in the late 19th century or later on by the modernist school of thought, that believed in being able to rationally engineer ideal urban structures in one big toss (see e.g. Hall 1996). One of the primal outcomes from this visionary advances usually comes in the form of a masterplan, a paper with arguably wide-ranging influences over people and space over many years to come - if implemented. Much has been written over the large master-planned projects of the last century, such as the cities of Chandigarh, Canberra or Brasilia, but also about smaller projects on the block or neighborhood scale. In most cases, the actual development diverged significantly from the vision and assumptions manifested in the plans, both physically and in terms of social appropriation of the built space (see Blake 1977, Kroll 2008). Nevertheless, most cities nowadays practice drafting masterplans, as well that envision their future development in structural and normative terms ("ecological", "sustainable "). Despite them being less impactful compared to new city developments (depending on the pace of growth or decay) they arguably still have significant, long-term implications.

This work is embedded in the argumentation line that sees cities as large complex systems and the discussion of how much planning and regulations are desirable for a successful development of towns on the one hand, and what planning should aim for on the other:
1) readily designing cities and quarters down to the shape of the roof or 2) support the self-organization process and actively plan with it while protecting public interests and the commons.

As of now, most urban masterplans are heavily driven by design and normative visions that lack the definition of measurable indicators and often ignore the principles of self-organizational forces and market mechanisms (Bertaud 2018). Thus, they are more likely to make wrong assumptions on people's behavior (e.g., in terms of location choices of developers, residents, etc.). This increases the chances of costly infrastructure being built at places that do not meet the demand adequately. A mismatch, in turn, comes with high costs - as once built, changes in the infrastructure are difficult to carry out (Oliveira 2016).
Therefore, some planners and practitioners advocate for the second approach mentioned above. Planning interventions are thus preferably launched in an on-demand fashion, responding to recent trends. This requires defining indicators that measure city statuses and their constant monitoring. For new developments, models with predictive abilities are required in order to help with the estimation of the impact of adding or changing elements within the city systems (Achary et al. 2017, Bertaud 2018, Batty 2017).

A vast body of literature exists on models and methodologies on correlating and predicting the behavior of city systems and people in relation to the physical built environment and socio-economic factors (Hillier 1984, Batty 2017, Stanilov 2010, also see chapter two) po-

tentially allowing to anticipate and simulate the impact of planning interventions early on during the design process.
Therefore, the adoption of these models into an accessible and integrated analysis framework might be a promising approach in coping with the issues described above. Especially in rapidly developing regions with high urbanization rates, where there is pressure to act fast, and a lack of institutional incentives to conduct in-depth testing of plans with detailed models is often present, this kind of toolbox might help to improve allocation of resources and increase the overall efficiency of urban projects and their impact in practice. But also for smaller cities and projects where resources lack for the employment of detailed impact studies, this kind of tools have the potential to positively contribute to the planning outcomes even more so if sufficient user-friendliness allows them to be integrated into an iterative planning and design process from the beginning. Despite the outlined probable contribution of modeling and simulation, in practice, these methodologies are still rather rarely applied, and (master) planning remains a highly design-centered exercise.

Aim of research

This work attempts to respond to the described issues and follows the call for a more data-driven, evidence-based approach to urban planning (Karimi 2014) for the use-cases depicted above by:

A) developing a methodology of a computational framework that integrates designing, plan evaluation, and monitoring of urban developments for employment in early design and planning stages to support evidence-based workflows.

B) exploratively applying the framework in the context of a new town project, aiming at the goal to find strategies for improving a plan's quality and resilience given uncertain growth projections.

A) *Developing the methodology - a computational framework:*

A prototype for an integrated, spatial analysis and simulation framework with the following characteristics:

1. Flexible application and high level usability for practitioners:
   - Parametric: - implementation in *Rhino/Grasshopper,* modular and expandable; mostly parametric, non-destructive workflow.
   - Flexible - few requirements towards data inputs, focus on basic spatial features, mostly retrievable via satellite imagery.
   - User-friendly - a user interface (UI) and standardized structure for manual inputs for sketching and comparatively analyzing scenarios.
   - Scalable - flexible altering of resolution and accuracy for calculations enable simulation on larger areas or increase the responsiveness.
2. A sketch-mode: allowing indicator guided sketching of design variants with near real-time feedback on selected performance indicators.
3. An evaluation-mode: calculation of all metrics and comparative statistics with functionality for interactive visualizations, filtering, and comparison of multiple scenarios.

In short: A toolbox to conduct spatial impact analyses of designs, while retaining a high level of accessibility and fast feedback (partly real-time analysis and feedback) to enable user or machine to engage in an iterative design and optimization process in early planning or design phases.

The technical backbone consists of a weighted, multi-modal graph model that is combined with an extended version of the Huff-model (see section 2.4.3), which stochastically distributes trips. It allows simulating interaction effects between land uses, associated socio-demographic properties, and points of interest through a multi-modal transportation network. Moreover, as the focus is laid on analyzing the configuration of basic spatial properties and related performance potentials. Even though the model only makes predictions in these fields it assists in improving the spatial configuration as one crucial puzzle piece in the equation of city planning. Due to the limited scope of the thesis, the model parameters and indicators cannot be validated on the ground for the case study. Instead, parameters will be set according to findings in existent literature. Moreover, the general goal of developing and applying the toolbox is to exemplify and to explore its application possibilities for several use-cases. Gained insights can then be used to implement more sophisticated indicator systems for future iterations of the model. Despite these limitations, it is hoped to contribute both in terms of outlying a framework worth to be further developed and gain some actual knowledge on the case of the Lanzhou New Area and new town developments (see below).

The theoretical and technical background of the toolbox is presented in chapter two, while its implementation, functionality, and user interface are explained in chapter three.

B) *Explorative application on a new town development:*

The prototyped framework is then applied, and exploratively tested on the Lanzhou New Area (兰州新区, LNA), PR China, with the following two use-cases:

1. Tracing the LNAs spatiotemporal development throughout four periods, analyzing and describing its evolution using the frameworks indicator system.
2. Testing of alternative, hypothetical spatial growth trajectories in comparison to the observed one and searching for more efficient ones.

The LNA is a large scale new town project in western China that targeted to reach one million habitants by 2030. However, throughout the last few years, it became more and more evident that this goal is far beyond reach, thus, rendering the scale and spatial allocation of the rapidly built infrastructures inefficient. Alongside with limited available data, the LNA is an interesting and relevant case study, both for testing the proposed toolbox and for engaging with the research question stated below. Detailed background information on the Lanzhou New Area is provided in chapter four.

Research objectives and questions- The objective of this thesis is to develop and showcase the prototype and to explore how such tools can assist in improving planning outcomes by closing the gap between top-down plans and actual urban developments. With the LNA as a case-study, the focus is narrowed to the issue of new town planning under uncertain growth projections. More precisely, the following research questions are examined:

- Do spatiotemporal configurations matter in regard to the efficiency of urban form measured over time?
- Which rules can be deduced to guide the design of spatiotemporal configurations that perform efficiently throughout all stages of the urban growth process?

Thesis structure. Following this introduction, chapter two ("Theoretical background and relevant literature“) positions the thesis in the nexus of urban modeling and simulation. One branch of contemporary planning practice critique is elaborated on, that argues for better integration of digital planning and simulation tools. It is followed by a review of relevant literature that focuses on the relation between the physical configuration of cities and their (socio-economic) functioning. Finally, the core models and methodologies that form the technical backbone of the toolbox are summarized with a focus on graph theory, the concept of accessibility, and gravity models.

Chapter three ("Design of the Toolbox“) reiterates the methodology, describing the implementation and functionality of the toolbox in detail; defining typical use-cases the framework is tailored for; describing the rationale behind the selected evaluation criteria, the technical implementation, and finally it’s applications and limitations.

In Chapter four ("Application: The Lanzhou New Area"), the toolbox will be applied to the case study Lanzhou New Area (LNA). First, details are given on data collection, parameter calibration, and the formulation of a performance indicator system. The spatial structure of the LNA is then evaluated at four points in time (2012, 2013, 2015, 2017) as well as at its finished state (as envisioned by the masterplan), using the proposed indicator system. Car-counts from satellite imagery are used as a proxy to incorporate levels of actual human activity into the analysis (next to physical structures). Results are then compared to alternative, hypothetical growth trajectories in terms of their performance within the indicator system and convergence patterns to the masterplan’s spatial structure.

Finally, chapter five ("Conclusion") reflects on the proposed research goals, questions, and the toolbox’s potential usability for planning and design, as well as its limitations and future research directions.

2. Theoretical Background and relevant literature

2.1 Overview

There exists a vast body of literature, models and methodologies on correlating and predicting the behavior of city systems and people in relation to the physical built environment and socio-economic factors with the potential to assist and complement the traditional approaches on urban planning and design (see for example Hillier et al. 1993, Batty 2017, Stanilov 2010, Gehl 2010, Zuend et al. 2016).
They range from static network-analysis over highly aggregated equilibrium- to detailed, agent-based models, simulating choices of generic individuals. As models, they assume a greatly simplified world that runs by a set of basic rules (f.e. profit-maximizing agents, perfect markets) while omitting many details of the real world situation.
Typically the more complex the model, the more detailed the results, the higher the number of parameters, calculation time, requirements toward data, and risk of overfitting. In many cases, data for complex models are not readily available and difficult to obtain. Moreover, they are more difficult to operationalize and interpret, and therefore pose high barriers for deployment in practice (see Batty 2008).
Despite less detailed results, simpler setups with few parameters have several advantages: they are less data demanding and easier to operationalize. Furthermore, a high aggregation level shifts the focus on the primal, general relationships inherent to (most) city systems, which allows a faster adoption of the model to different contexts and cases. Here, results can rather be interpreted as probabilities or natural potentials attributed to urban structures.[1] In case of new developments and especially new town projects, being able to estimate the configuration of potentials in relation to the planning proposals can be particularly useful to create designs in greater harmony with the self-organizing forces of settlements - in the spirit of the second approach on urban planning outlined above.

2.2 Critique on contemporary planning practice

When looking at cities as complex systems, urban masterplans with time horizons of often 20 years intrinsically (at least partially) struggle to provide an adequate framework to manage the ever ongoing change processes in cities.
This is especially apparent in new town projects where the whole town (or its skeleton) is designed at once. Planners face many uncertainties, hence basing decisions on bold assumptions that might turn out as incomplete or wrong over the years to come (see Karim et al. 2014, Yu et al. 2013b).

Moreover, as Bertaud (2018) notes, most masterplans are heavily driven by design and normative visions that lack the definition of measurable indicators and often ignore the principles of market mechanisms. Thus, they are more likely to make wrong assumptions on people's behaviors in terms of location choices of developers, residents, commuting. This increases the chances of costly infrastructure being built at places that do not meet the demand adequately. A mismatch, in turn, comes with high costs - as once built, changes in the infrastructure are difficult to carry out (Oliveira 2016).

---

[1] For instance, a high potential for the commercial center of a city would probably be computed for areas nearby highly integrated streets)

According to Ye and van Nes (2013b), urban plans and the spatial structure they propose to define the cornerstones of a city's future development both physically and socio-economically. In their view, the more the spatial structure matches the social logic of space (Hillier et al. 1984) in the first place, the faster and more significant the chance new towns or other developments reaches their full development potential and provides an attractive environment to live in (as one factor among others).

The concept "social logic of space" is based on the assumption that physical elements, second to none the street network, affect and mediate the interaction of all the other aspects of the urban system (such as behavior and movement patterns of people and thereby also the distribution of land uses and activities) - with the peculiarity of being long-lasting and comparatively easy to quantify.

However, as quite some studies pointed out (e.g., Karimi et al. 2014 with an example on English new towns or Bertaud (2018) claiming that most masterplans don't pay sufficient attention to the way markets work and affect the functioning and form of cities) it occurs that the development goals of masterplans are often not coherent with the spatial-structure they propose. In other words, the top-down development contradicts the market-level logic or, put differently, the self-organizing nature of city systems.

Obviously, in either case, organically grown or top-down designed cities will somehow function. However, as Karimi (2014, p13) put it: "[planned] Centers are attractors in the urban grid, but if a center is an embedded attractor, it will have additional potentials to act as an economic focus in the area", thus, arguing that, if centrally allocated, land uses are best assigned in coherence with the nature of the towns spatial structure.

In this context, Ye (2014) observed an "urban maturation process" in dutch new towns, which is the process of the spatial structure of a town moving towards an equilibrium-like state with space's social logic. The cities habitants and stakeholders mainly drive this process in a bottom-up fashion where faster-changing aspects such as land-uses or buildings are adapting to slower changing elements such as the streets. Al Sayed et al. (2016) noted a similar process when looking at time series data for Manhattan and Barcelona, two cases with arguably strong planning regulations in place. The time and costs this process consumes are highly depended on the initial spatial structure and policies in place, thus anticipated these adoption processes appears to be beneficial.

Furthermore, Bertaud (2018) argues for incorporating (at least) basic urban economic models into planning practice to map development pressure and monitor shifting equilibriums in demand and supply. Anticipating these market forces would help to conclude a more realistic assumption during plan drafting of, for instance, large infrastructure projects, and thereby lowers the risk of a mismatch between infrastructure investments and actual developments

In summary, the above-mentioned authors argue for the application of models and simulation to predict and simulate the (potential) impact of interventions and the urban system's reaction to them, in order to arrive at overall better solutions and ensure coherence with the social logic of space. Besides the demand for a quantitative indicator and monitoring system, one could summarize their critique, in Karimi's words, as a call for "evidence-based planning“ (Karimi et al. 2014).

2.3 Urban form and performance: relevant literature

The defining nature of cities can be thought of as places with high densities of interaction and exchange potentials (Batty 2017). The high level of potentials is enabled by people living and working together in high proximity. Density derives itself from short travel time (and costs) from a place to all people, jobs, and locations of interests. Travel times and the distribution of movement flows are closely related to the configuration of the road and transportation network. In turn, land uses such as businesses, shops, and services choose their location in response to this configuration as they rely on by-passers and access to as many customers as possible.
In a similar way, systematic relationships between urban form and people's aggregate behavior have been found, allowing for, sometimes more sometimes less accurate, simulation of idem. Therefore, certain impacts of different spatial arrangements on human behavior can be estimated in advanced and utilized in the planning process.

The following part is a brief review of literature on the matter of correlating spatial properties of cities with socio-economic phenomena and serves to define the indicators to be computed by the toolbox (see section 3.2). The papers that were regarded most significant are overviewed in Table 2. They were chosen according to the earlier outlined goals of the toolbox and its requirements toward data inputs and its level of detail. Thus, the search focused on studies correlating essential urban spatial structures (road network, built volume, land uses) with their impact on the functioning of a city.

Marcus (2007) introduces the term spatial capital, which is defined by three equally weighed indicators corresponding to the one mentioned above (see table for details). All analysis is being conducted using a 500-meter radius in network distance (as this is a typical walking distance). After all, all values are positively correlated with spatial capital, which can be interpreted as the potential for "urbanity“[2] and high levels of interactions.
Despite solely focusing on the street network, Hillier (2009) extends his analysis from computing indicators to the additional dimension of their spatial distribution. He proposes the meta-indicator of spatial sustainability, composed of the sub-categories ecological, social, and economical. He conducts SpaceSyntax analysis for multiple radii and correlates the spatial structure of the street network and resultant centrality values with tax bases, the distribution of center locations, and crime rates to prove the predictive capabilities of network analysis on the three indicators (see table 2 for details). Taking the, rather organically evolved, city of London as a case, he further shows that naturally grown cities score higher in spatial sustainability compared to many centrally planned projects as planners tend to overlook the social-logic inherent to spatial structure.
Ye et al. (2013a), conduct similar analysis as Hillier but add two additional factors to their comparative analysis of the spatial structure of organically grown and new towns: 1) A measure of density, where they use the space matrix methodology to cluster housing into diffe-

[2] Marcus 2007 (p. 3), defined urbanity as follows: "*urbanity, both socially and spatially, is primarily constituted by high accessibility and high diversity*“. In this sense the term tries to capture what the Author regards the very nature of cities, that is people moving to cities in order to be in proximity to other people and opportunities.

rent categories of "urbanity“ using the two indicators FSI (Floor space index) and GSI (ground space index). 2) a measure for land use mix (termed MXI, mix-use index), which computes the degree of diversity between the three land use categories work, residential and commercial. The values are aggregated into grid cells in order to make them relatable to one another. Furthermore, the authors used a natural break clustering algorithm to separate the values of each indicator into high, medium, and low. These values are then used to compute the degree of urbanity for each cell. They conclude that, compared to organically evolved settlements, new towns tend having significantly less highly urban areas that are furthermore less correlated with the areas of high accessibility (as measured by network analysis) naturally resulting in worse overall accessibility to centers. In a follow-up study (Ye et al. 2013b), they use the same indicator system to analyze and compare several new-town projects in the Netherlands (and China in Ye et al. 2014) toward their degree of urban maturity. Urban maturity is defined as the match-rate of highly urban areas (represented by grid cells) with areas with high centrality values (Closeness and Betweenness for local and global radii). Additionally, they cross-checked their results with socio-economic indicators, where a higher degree of urban maturation was associated with a better performance in the socio-economic indicators. In a more recent study, Karimi et al. (2014) returned to the topic of new-towns as well and came to similar results as Ye et al. (2013b): Naturally grown towns tend to perform better compared to their central planned counter-parts. He concludes that insights of network analysis and simulations should be incorporated in the design process early on (see table 1 for details).

The urban performance chart by the MIT CityScience group (see Zhang 2017), consists of 17 indicators subsumed in four categories to describe the performance on the neigh-

Indicator	Organically grown tows	New towns
Network and accessibility	Centers are clearly defined where local and global centralities overlap — organically distributed sub-centers. Neighborhoods typically do have fuzzy borders to one another, allowing them to interact and contribute to the formation of local centers. Good and distributed walkable access to centers	The most integrated streets are usually motorways and intersections. That, thus act as local separators instead of places of interaction. Global and local centrality values are less correlated, decreasing the potential for vital centers to emerge. Residential neighborhoods are prevalently designed „compounds“ with inward facing local road systems, mostly connected by a „vehicular super-grid“ to other areas of the town.
Movement patterns	Pedestrian flows are highest in the town centers and decrease with distance to them with local peaks in multiply centers of lower hierarchy.	A combination of a low degree of walkability and land-use mix lead to a very low level of pedestrian movements in most areas of the town. Instead, only a few strong peaks in designated centers such as malls emerge. This leads to lower degrees of perceived safety in many areas and hinders social cohesion and exchange.
Land use distribution	Land uses usually follow centrality patters with commercial and activity intensive uses locate in the most integrated and accessible areas. Further, activities extend along major (highly integrated) roads throughout the city.	The degree of land use mix is much lower compared to old towns. Land use adherent to centers (e.g. Commercial) is not matched with highly integrated regions, thus presenting a mismatch and lead to longer commutes and worse accessibility.
Block size	In central regions, the bust structure or blocks tend to be more, allowing for better inter-connectivity and mix use. In residential areas, block sizes tend to be relatively larger, which leads to more concentrated pedestrian flows concentrated along a few roads.	Block size in most new towns is rather large, also in areas that are supposed to act as centers. This decreases intra-neighborhood accessibility and presents an obstacle to navigation

Table 1: Spatial characteristics of organically grown towns and new towns. Source: Author, based on Karimi 2014.

Authors / main concept	Key indicators (depended varia-	Independent Variables and calculation method	Discription
Hillier 2009 *Spatial sustainability*	Economic sustainability	- correlation of accessibility with values on tax bands and the location of centers	The author discusses the term spatial sustainability, that is defined by three sub-categories: ecological, social, and economic and relates the term with Space Syntax analysis (betweenness and closeness). The latter is conducted for multiple radii and correlated with tax bases, distribution of center locations, and crime rates to prove the predictive capabilities of network analysis on the proposed indicators in organically grown cities.
	Ecologic Sustainability	- Travel distances, minimized if cities develop along „natural“ trajectories, leading to more evenly distributed centers with good accessibility.	
	Social Sustainability	- Community formation (accessibility and integration within and to surrounding neighborhoods) - Security (based on neighborhood internal pedestrian flow distribution)	
Marcus 2007 *Spatial capital*	Accessibility/ Centrality	Local accessibility (betweenness) as calculated with SpaceSyntax methodology	Marcus Introduces the term spatial capital as defined by the three, equally weighted, indicators on density, accessibility, and diversity. A high spatial capital is associated with acting as a busy center for social and economic interaction and exchange.
	Access to density	Sum of floorspace within 500 meters via network distance	
	Access to diversity	Number of plots within 500 meters via network distance	
Bielik et. al. 2018 *Walkability*	Walkability potential	Access to areas with high centrality (SpaceSyntax) values, calculated by applying a distance decay function.	This study concludes that, with reasonable accuracy, the location of walking attractors and thus the walk-score of an area can be estimated just using the street network as input. For the case study (Weimar) exclusively using the network as an input, 85% of the walk-score results could be explained.
Ye et al. 2013a *Spatial flaws of new towns*	Accessibility / Centrality	Local and global SpaceSyntax accessibility analysis, *a high interconnectivity and wide distribution of high centralizes is regarded as positive.*	The authors compare organically grown cities with new towns and point out spatial properties explaining the better performance of many old towns. The higher the indicators, the higher the urban intensity or degree of urbanity, which is generally regarded as a positive attribute by the authors.
	Density and building typology (Space matrix)	Clustering building typologies between suburban and highly urban by the variables floor space index and ground space index	
	Diversity (Mix-use index)	Measures mix of three land use categories (residential, work, commercial) within an area.	
Ye et al. 2013b *Urban maturity*	Match-rate	Refers to the match-rate of areas classified as highly urban (based on built-up density and use mix) and areas with high accessibility values (space syntax).	The authors introduce the concept of an urban maturation process, that can most prominently be observed in new towns. In mature city functions and uses are in coherence with its social logic of space, mainly defined by the towns transportation network. The degree of maturation can be approximated with the suggested match rate. A high match-rate is positive, as functions are more effectively distributed, naturally balancing demand, supply, and accessibility.
Long 2017 *Urban economic vitality*	Economic vitality	Most relevant variables in explaining economic vitality (represented by user activity on the online platform dianping): Intersection density; # of public amenities; distance to closest transit station; use mix structure; pop. density	Long conducted a country-wide study relying on open data. The analysis operates on a 1x1 km grid and correlated several physical properties with spatialize data on economic activity.
Rode et al. 2015 *Transportation and urban form*	Infrastructure and operational costs	Relation of population density and infrastructure investment (sewage, water, road networks) per capita	A literature review on studies analyzing the impact of the spatial structure of towns in climate, health, equality, and mobility-related dimensions. Most reviewed studies operate on a highly aggregated level (complete cites as spatial units). Cities from all over the world are systematically compared eg. by their built-up density, population distribution, CO2 emissions, accident rates acne economic productivity. Finally, significant relations were found between density levels and described indicators. These results can be used to position cases or new town projects in comparison and estimate its performance trajectory.
	Health and ecology	Km traveled by transport modes m times avg. accident rate per km and CO2 emissions per km traveled.	
Bertaud 2018 *Order without design*	Size of the effective labour	Average percentages of total jobs accessible in less then one hour	Bertaud argues for better integration of urban economics with planning practice. He sees cities primarily as labor markets that benefit from economies of agglomeration. Thus, the primary responsibility of planners is to ensure mobility (access to jobs) and affordable housing prices. Moreover, he holds the view that planning should follow an on-demand fashion intervening if necessary, which in turn requires a monitoring system for identified key indicators while taking the self-organizing forces (or market forces) seriously. Therefore, indicators should be extended by a temporal dimension.
	Avg. commuting time to jobs for different modes	Overall average commuting time to jobs weighted by population (depends on network, population and job distribution)	
	Changes share of formal and informal housing stock	Indicates how well the demand for housing is accommodated in a city. High levels of informal growth should lead to a review of the current planning strategy.	
	Floor space area consumption per capita	In the mid and long run the changes in this indicator can be can give insights on the balance between supply and demand on housing. Sinking values indicate low demand and potential over-supply of housing.	
	Attractiveness of transit vs. car commutes	Indicator compares the door-to-door time between the two transportation modes.	
	Cohesion to market market forces	Relation of property/housing prices to distance to town center and its change over time	
Zhang 2017 *MIT urban performance chart*	Diversity	Calculates the shannon index of diversity by with the classes residential, employment, third places and cultural. The index peaks when all classes are represented with an equal share.	17 Indicators subsumed in 4 categories to describe the performance on the neighborhood level. Next, to spatial properties, socio-economic data such as social milieus of residents are considered. High performance is understood as an energy efficient, vibrant, and divers quarter in the sense of the „creative city“ and entrepreneurial city.
	Proximity	Measures access to parks and public transport stations as well as the intersection density.	
	Density	Measures density of residence, employment and POI density (third places, cultural places)	
	Energy	Measures energy efficiency, in terms of output and inputs for buildings and transportation.	

Table 2: Literature on urban form and performance. Source: Assembled by Author

borhood level. Next, to spatial properties, socio-economic data such as social milieus of residents are considered. High performance is understood as energy efficient, vibrant, and a socially diverse quarter in the sense of facilitating the creative and entrepreneurial city. Another feature is the use of an exponential decay factor (in the fashion of gravity functions) for accessibility-related measures (e.g. access to parks and amenities). An issue in the context of this thesis is the higher requirements for data (e.g. socio-economic ones).

The studies of Long (2017), Rode et al. (2015), and Bertaud (2018) looked at the relation of urban form and function from a more aggregated perspective.
Long (2017) conducted a countrywide (China) analysis of urban areas that operate on the resolution of a 1x1 km grid. He correlated several physical properties with spatial data on economic activity represented by user activity and listings on the online platform dianping (similar to yelp or google places). The three most relevant variables in explaining the variation in economic activity were the intersection density, and the number of public amenities followed by the distance to the closest transit station (see table 2 for more details).
The meta-study "Accessibility in cities: transportation and urban form“ by Rode et al. (2015), focuses on differences in associated costs between cities with varying built-up and population densities as well as the degree of urban sprawl. Costs are estimated for infrastructure investment per capita, mobility, modal choice, and the use of private cars, indirect effects on health and the environment (via pollution and fatal accidents related to cars), social equality, and the economy. Their analysis of cities worldwide shows that denser cities are correlated with significantly lower per capita costs in most of the areas investigated (for instance, see Fig. 1). Moreover, denser cities enable efficient functioning of mass transit, which in turn improves health, and equality by reducing the probability of private car usage and the provision from affordable mobility.
Bertaud (2018) primary regards cities as labor markets, that attract more people as they provide job opportunities and attract more companies as they provide a large pool of labor and talent. The central role of planners, therefore, is to ensure an efficient transport system and affordability for households and firms or put differently maximize the centripetal- while minimizing centrifugal-forces occurring in city systems. Therefore, among the most crucial indicators defining a city is the effective size of its labor market and the accessibility to jobs. A cohesive labor market is here defined as all jobs within a travel radius of one hour. These measures are, in turn, strongly related to the towns transportation system and the distribution of both housing and jobs. Factors, all of which can be summarized with indicators on the average commuting time to jobs and the average number of jobs within one hour travel time. Also taking a macroeconomic perspective, Graham et al. (2010) investigate the dis-

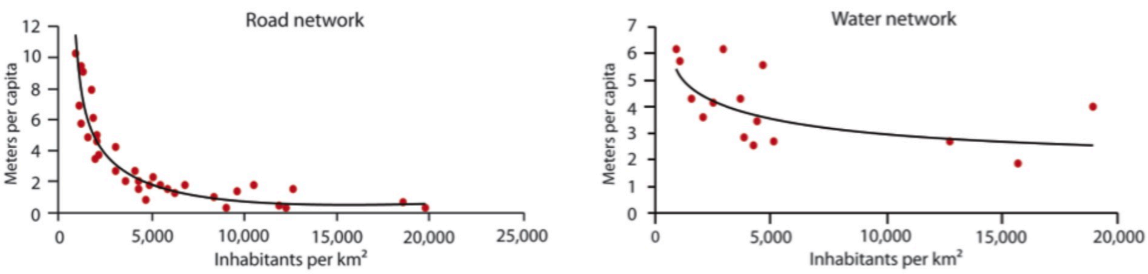


Fig. 1: Impact of urban density on road and water infrastructure costs. Source: Rode et al. 2015.

tance decay moderating economies of agglomeration[3] and empirically estimate the decay parameter for several sectors, and the overall economy (the authors used a basic gravity function with one parameter, see section *gravity function* below). Therefore, the degree of agglomeration depends both on the spatial distribution of land uses and the transportation network. With the estimated decay factor, it is possible to evaluate, after which distance two localities stop (significantly) contributing to each other's degree of agglomeration. Finally, the authors estimate that an increase of the agglomeration degree by 10% is associated with an increase of total factor productivity by 0.4% in average and by 0.8% for service-related industries, thus being an important measure for capturing economic potential.

The general tenor of the reviewed papers can be summarized as follows: 1) The physical structure of cities is one influential impact factor on the socio-economic configuration and functioning of cities 2) Some of the impacts can be simulated and predicted (sometimes with higher, sometimes with lesser confidence) which leads to the conclusion that 3) It is desirable to integrate simulation measures into the planning process, which is, however, seldom the case in practice.

2.4 Urban modeling

Technically and conceptually the tool-box is rooted within several modeling techniques and underlying theories that will
briefly be examined in this section.

2.4.1 Graph theory

As famously stated by Tobler (1970,) the most fundamental principle in geospatial analysis is the cost of overcoming space or simply distance. Distance mediates the interaction and relation between two places of interest. There are different ways of measuring distance. In the spatial context, the most straightforward way to do so is using the euclidian distance, which is the as-the-crow-flies distance measured in meters. More sophisticated approaches try to capture the cost of overcoming space in monetary, physical, psychological, and time-related terms. This higher complexity in relating places can be represented and modeled using graphs. A graph con-

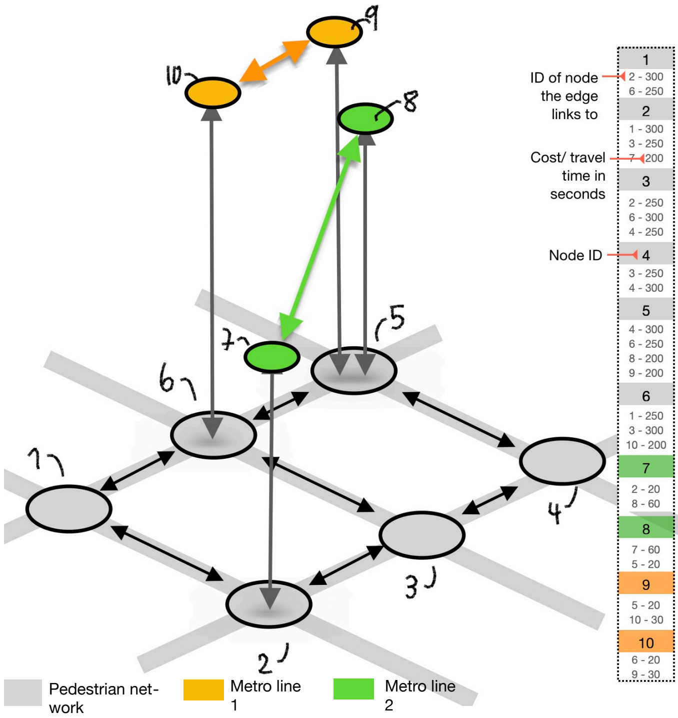


Fig. 2: Multimodal graph structure for pedestrian and public transit segments. Nodes are represented by the circles, edges by the arrows. Table on the left illustrates the data structure of the graph in form of a data tree. Source: Author

[3] In economics agglomeration effects are regarded as externalities related to the spatial proximity of economic activity or activities in general. Empirically there is much evidence that higher levels of agglomeration are correlated with higher productivity (see, eg. Melo et al. 2009). Reasons for these positive effects arise from larger talent pool firms have access to, quicker and cheaper exchange of services goods and ideas and more efficient co-use of public services and goods (see Graham et al. 2010).

sists of vertices (or nodes) and edges. In the context of modeling cities, vertices can be interpreted as locations of interest such as shops, squares, blocks, or street segments. Edges act as the connection between nodes. They are either directed (allowing movement only into one direction) or undirected (allowing bidirectional flows). Each edge can be attributed with weights representing the cost of passing it (e.g., the travel time from one node to another). Fig. 2 presents an example of a transportation network represented by a graph.

A well-known approach to model the psychological cost of overcoming and navigating through space was introduced by Hillier (1986) as SpaceSyntax and has since excessively been tested empirically. SpaceSyntax proved to have high explanation capabilities for pedestrian movement frequencies. In many cases, 60 to 80% of pedestrian flows can be explained by solely analyzing the physical configuration of the street network, without considering other aspects of the built environment nor socio-demographics (Hillier et al. 1993). Instead of the metric distance between two locations (nodes) Hillier introduced the angular-distance which (ideally) requires to 1) using Axial lines (lines-of-sight) instead of road center-lines to model the urban network and 2) measures the difference in direction (in angles) when moving from one node to the next (see Hillier 1996).

2.4.2 Network analysis

Graphs are ideally suited to study all kinds of networks (e.g. social networks), including street or transportation networks. That, in combination with shortest-path algorithms[4], open a wide range of application for analysis and simulation. The shortest path is defined as the route between an origin and destination node with the lowest value after summing up the weights of all edges crossed on the route. If the edge weights are estimated accurately, the shortest path calculations can be used for reenacting peoples route-choices. Given a known distribution of origins and destinations (see below), this allows for the prediction of movement flows and volumes through the urban network.
Besides, several centrality measures can be computed for graphs and the networks they represent. Most widely used within the realm of street network analysis are Closeness and Betweenness centrality which are calculated as follows:
Closeness centrality, C for a node i is usually defined by:

$$C_i = \frac{1}{\sum_j d_{ij}} \qquad (1)$$

Where $d_{i,j}$ is the distance between node i and node j. Additionally a maximal search distance can be defined, where only nodes with $d_{i,j} <= dist_{max}$ are used for the calculation. A high closeness centrality value of a node means, that all other nodes can be reached comparatively fast and vic versa that it can be reached comparatively fast from all other nodes in the network (see Newman 2010). In the context of street networks, segments with high closeness centrality are usually associated with city centers (Hillier 1984). Moreover, by changing the maximum distance of for analyis, the likely location of local and city-wide centers can be computed.

[4] A widely used shortest-path algorithm is the Dijkstra algorithm, see Kairanbay (2013).

Betweenness centrality can be expressed by:

$$B_k = \sum \frac{sp_{ij}(k)}{sp_{ij}} \quad (2)$$

Betweenness measures how often a node k was passed ($sp_{ij}(k)$) on the shortest paths (sp_{ij}), connecting each node with all other nodes in the network (see Newman 2010). Nodes with high betweenness values can be interpreted as important connectors or "bridges“ within a network. Empirical studies of street networks show a high correlation of betweenness, and the location choice of entities depending on footfall, such as shops and other commercial facilities. As outlined above, a maximum distance can be set when calculating betweenness centrality, shifting its interpretation to different scales.

Fig. 3, visualizes both betweenness and closeness centrality for the medium sized city Chartak (Uzbekistan) for a maximum distance of 1600 meter and 800 meter. Note, how both centrality measures are peaking at the location of the bazaar and the train station.
Besides, centrality measures proved to have a high predictive capability for a wide range of other cases: For instance, Bielik et al. (2018), approximated walk-scores[5] for the city of

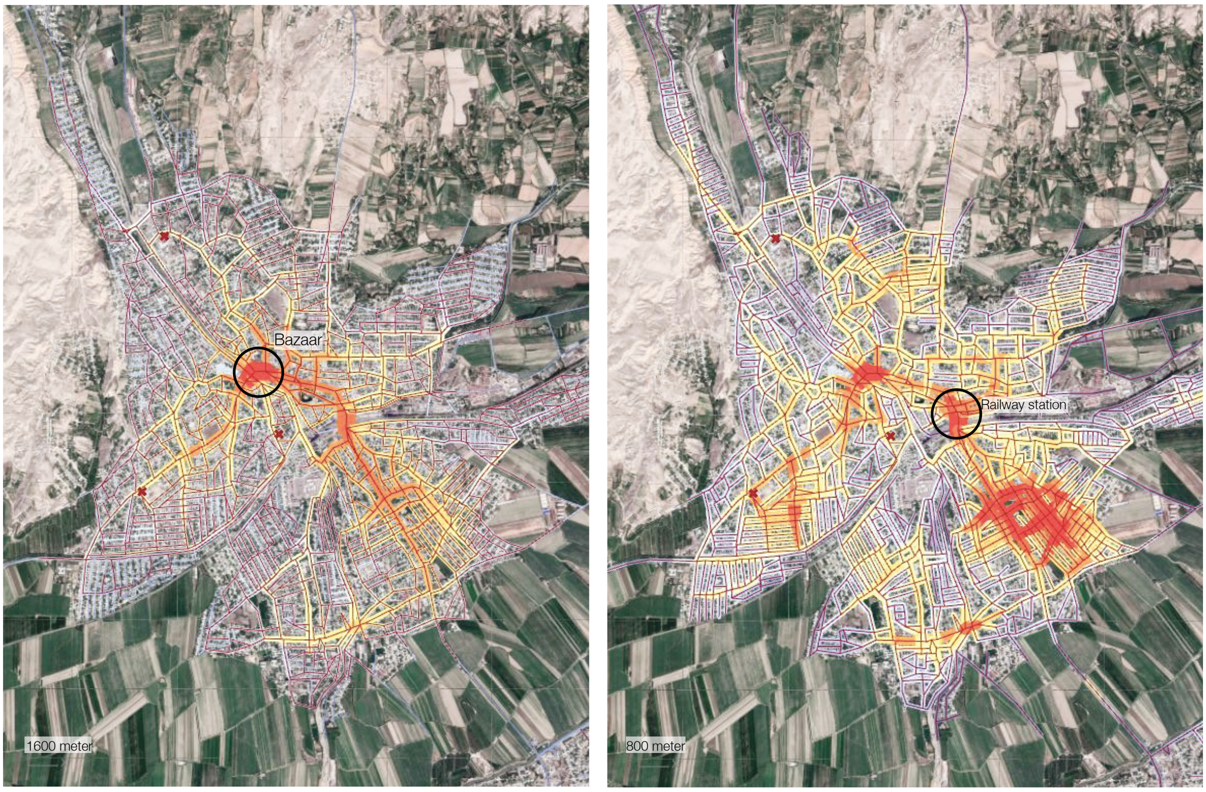


Fig. 3: Betweenness and Closeness centralities for the city of Kagan. High values coincident with high densities of commercial facilities. Left: analysis radius = 1600 meter ; Right: analysis radius = 800 meter. Betweenness is depicted by thickness, closeness by color (red = high, blue = low). Source: Author

Weimar (adjusted r-square of 0.8) by computing accessibility to highly integrated street segments (street segments with high betweenness centrality were used as proxies for wal-

[5] The walk score is a measure for walkability. For a given places is calculated by measuring the access to a set of amenities. Each amenity within a maximum distance of 30 min adds to the final score while it is discounted for distance and similar amenities nearby (f.e. adding a element restaurant to an area will not improve the sub-scoring on restaurants; neither does a restaurant that is almost or more then 30min away from a location). High walk score had been significantly correlated with property prices and health among others (see Oishi et al. 2015 and walkscore.com for a description of its methodology).

king attractors). The only necessary data input was street-centerlines. In their quest to estimate the spatial distribution of jobs in the cities of Baton Rouge and Barcelona, Wang et al. (2011) and Porta et al. (2012) found closeness and betweenness centrality values as the most relevant indicators as they alone explained between 30 and 80% of the spatial distribution of jobs.
The high relevancy of the two centrality measures lies in their nature of 1) predicting footfall potential (betweenness), which is an driving factor for the location choice of businesses that depend on by-passers and 2) expressing the degree of accessibility or how easy one place can be reached from all other places in the case of closeness. Which is arguably important e.g. for the location choice of businesses and offices, as they benefit from access to the labour force and costumers/partners. In this context Batty notes: "*locations are built on interactions while interactions are canalized by transportation networks, that structure movement and connect the locations"* (Batty 2017, p53).

For both presented centrality measures it is possible to incorporate weights into the calculation. In the case of urban analysis, it typically makes sense to set weights using the number of people living in an area adjacent to a node or the amount of floor space, jobs etc.. For Instance, a betweenness weighted with population arguably helps to get closer in estimating the pedestrian frequency within an area or footfall. It can also make sense to further mediate the weights by applying a distance decay factor (in a gravity function fashion, see below) to them, to model the issue of people being likely to visit near-by locations more frequently than locations far off (see Ewing et al. 2010).

2.4.3 Gravity models

Gravity models are among the most frequently used model category within geography. Originally introduced by Hansen (1959), they effectively operationalize Toblers first law of geography that is "everything is related to everything else, but near things are more related than distant things." (Tobler 1970, p. 234). The general gravity function is defined as:

$$G_i = \sum \frac{W_j^{\alpha}}{e^{\beta * d_{ij}}} \quad (3)$$

Where G_i is the gravity induced by destination j to origin i. W_j is the weight of destination j (e.g. the total floor area of an location), while α is a parameter that scales the effect of the weights. e is the natural logarithm. d_{ij} is the distance between origin i and destination j, while β is the distance decay parameter. The parameter α and β are ideally estimated empirically (see Sevtsuk et al. (2017) for an applied example of the parameter calibration)**.** While in practice α is often left to 1, the β parameter is crucial. Multiply empirical studies concluded that, in the urban context, a β-value of 0.002 (when distances are measured in meters) generally delivers good results for pedestrian movements. However, the β-value naturally differs from destination type to destination type (e.g. subway-station, restaurants etc.), from city to city, and transport mode to transport mode. See Thomas et al. (2003) for an overview, who conducted an in depth study of the variation of the beta parameter for the case of Belgium.

The basic gravity model can be extended to model competition between places (such as retail centers or parks) and peoples choices on which place to visit. This variation is also known as the Huff-model (see Huff 1963). Its basic assumption is that the probability of a potential visit of a person to a point of interest is a function of the point's attractiveness and its distance to the person (basically equals G_i) relative to the gravity received by all other points of interests. Thus, the function can be defined as:

$$Visitors_i = \frac{\frac{W_j^\alpha}{e^{\beta * d_{ij}}}}{\sum \frac{W_j^\alpha}{e^{\beta * d_{ij}}}} * pop_j \qquad (4)$$

Where Visitors i is the number of total visitors to a point of interest**.** The first term represents the relative gravity (gravity between point of interest i and origin j divided by the sum of the gravitation between origin i and all points of interest). The second term, pop j, is an additional weight attributed to origin j, resembling the population or potential costumers.
This basic version of the model, distributes the whole population among all points of interest. However, one could argue, that if a origin is very far away from any point of interest, no visitor at all will originate from that origin. This can simply be modeled by adding another distance decay parameter to the model, affecting the weighting parameter pop j, resulting in:

$$Visitors_i = \frac{\frac{W_j^\alpha}{e^{\beta * d_{ij}}}}{\sum \frac{W_j^\alpha}{e^{\beta * d_{ij}}}} * pop_j \frac{1}{e^{\beta 2 * d_{ij}}} \qquad (5)$$

Where β_2 is the additional distance decay parameter. This approach can be, and is often applied in modeling origin-destination matrices that form the basis for traffic and footfall modeling (for instance see Hensher 2004**)**.
To capture the effect of economies of scale, a kernel-density function could be incorporated. This would effectively replace the weight of an individual point of interest (POI) with the (distance discounted) weights of all (similar) POI within a designated radius.

2..4 Accessibility

Accessibility is a widely used concept for evaluating and comparing the performance of (transportation) infrastructure. As Piovani et al. (2018) put it: "*Accessibility associates some measure of opportunity at a place with the cost of actually realizing that opportunity*" (p.10). Multiple variants for modeling accessibility exist (Geurs 2004). The accessibility indicator chooses here is simple and easy to interpret:

$$A_i = \frac{1}{n} * \sum \frac{E_j}{c_{ij}} \qquad (6)$$

Where A is the accessibility indicator for area i, E is the measure of opportunity (e.g. number of jobs) at place j and c_ij the cost or travel time from area i to j. N is the total number of areas. In contrast to the gravity function, distance discounts the opportunity linearly. The metric is especially interesting for comparing the situation before and after changes to infrastructure and transportation networks.

3. Design of the toolbox

3.1 Design goals, scope and model structure

To follow the call for a more "evidence-based“ approach on urban planning that pays more attention to the self-organizational forces in inherent to cities, the goal is to develop an integrated toolbox that operates with data obtainable from typical urban masterplans and from satellite imagery.

On the one hand, this allows for quick analysis of masterplan designs without any additional data needed, while on the other hand, the focus on satellite imagery has the advantage of relying on a homogeneous, globally available data source that can hardly be manipulated. Given the recent advancements in computational object recognition, most basic urban features such as footprint, FSA, land uses, street networks, and general data on traffic volumes will soon be retainable with decreasing effort from those pictures with sufficient accuracy (see section limitation and outlook below for an overview).
Another advantage lies in the possibility of frequently updating the data at relatively low costs. This makes it an interesting source for monitoring urban development.

The general use-case of the toolbox is thought of for projects where larger, sophisticated models and simulation frameworks are not applicable due to data, budget, or know-how constraints. It offers simplified analysis for planning and design cases that else wise would be executed without any quantitative reasoning.

Therefore, in terms of complexity, the toolbox aims for a rather high level of aggregation with few parameters, to ensure greater transparency, better transferability to other cases and lower requirements toward input data and user expertise.

The presented scope mainly allows for network and accessibility-related analysis in interaction with basic land use categories (e.g., residential, commercial) and points of interest (e.g., parks, shopping centers, transit stations).
However, even this relatively limited data already enables to compute a wide range of indicators for evaluation and monitoring, as will be shown in the section below. Besides, the implementation in the Rhino/Grasshopper environment, as well as the modular structure, enables for rapid extension and adoption of the model to new requirements and cases.

The model is aimed to provide the following characteristics and functions:

1. Flexible application and high level of usability for practitioners:
   - Parametric: implementation in *Rhino/Grasshopper,* modular and expandable; mostly parametric, non-destructive workflow
   - Flexible: few requirements toward data inputs, focus on basic spatial features, mostly retrievable via satellite imagery
   - User-friendly: an user interface (UI) and standardized structure for manual inputs for sketching and comparatively analyzing scenarios
   - Scalable: flexible altering of resolution and accuracy for calculations enable simulation on larger areas or increase the responsiveness

2. A sketch-mode: allowing indicator guided sketching of design variants with near real-time feedback on selected performance indicators
3. A evaluation-mode: full calculation of all parameters and comparative statistics with tools for interactive visualizations, filtering, and comparison

The technological core of the framework is a (weighted) graph-model of an area's transportation network (streets, transit, etc.). Optional weights are population-, job- or POI density. POIs can alternately be placed manually. All network related calculations are performed within Grasshopper using the plugin SpiderWeb and its implementation of the Dijkstra shortest path algorithm.[6] For the full calculation of betweenness and closeness values components of the DecodingSpaces plugin are employed, the user interface was implemented with the library Human UI.

Basic data inputs:

- required: street network (line/polyline)
- optional: transit network (polylines)
- optional: population and job distribution (geocoded as .shp/csv) or estimates from bitmaps (e.g. LU plans); building footprints)
- optional: points of interest (such as transit stations, pharmacies, restaurants etc.)
- optional: satellite images as background

Basic parameters to set:

- avg. travel speed per road category
- avg. travel speed of transport mode
- avg. waiting time at stations per transit mode
- Parameters for the gravity function
- Optional: additional attributes for land uses (e.g. Population sub-groups, energy consumption etc.)

All computed indicators can be re-mapped to a uniform analysis grid, allowing for comparison and filtering across all computed measures and scenarios. Next to analysis in Rhino/Grasshopper, results can be exported as shapefile containing the grid points and network segments with all data points attributed to them. Several pre-calculated scenarios can be loaded and compared via a graphical user interface displaying controls and diagrams, while spatialize results are directly visualized in the Rhino canvas.

Related software and tools

There exists a wide range of spatial analysis software and plug-ins that, the one way or another, overlap with the functionality of the proposed toolbox. For instance, ArcGIS offers a wide range of sophisticated models and analysis possibilities both for network and graph related applications, among many more. The often-cited UNA toolbox (Sevtsuk et al. 2017) (available as a plug-in for ArcGIS and Rhino) also focuses on network analysis in the urban realm and is technically based on a weighted graph-gravity model. However, it does not offer the possibility to model transit networks or to analyze calculated data on the fly in the same environment without switching software and changing data formats. Moreover, the

[6] See Kairanbay (2013).

spatial resolution can not easily be scaled as proposed by this toolbox (by using a low-res grid and underlying distance matrix). The graphical programming interface inherent to Grasshopper and the modular setup of the toolboxes interface (see below) enables us to add and integrate project-specific indicators quickly into the workflow and the analysis. Another advantage of a full implementation within the Grasshopper environment is the high degree of parametrically, which enables a non-destructive workflow. In case a change is made to, say, the base network, a parametric setup updates all necessary parts of the definitions, including computing indicators without further actions needed to be taken by the user, which is the case when sub-results (e.g., number of population, visualization of results) are processed in unconnected workflows of different software.
Besides a range of similar models implement in Grasshopper emerged over the last years, however most of them are more specialized and, more importantly, operate on smaller scales such as block or neighborhoods levels (see for example Dogan et al. 2018, Herthogs et al. 2018, Wilson et al. 2019) whereas, this toolbox aims to cover whole cities. Finally, with the sketch-mode, this toolbox offers a streamlined and integrated functionality to sketch-out new scenarios with fast indicator driven feedback.

3.2 Evaluation indicators

Build on the possibilities opened up by a weighted graph-gravity model, an indicator system was developed that is rooted in the reviewed literature and the proposed requirements toward planning, as discussed by authors in section two.

Basic indicators:

Table 3 outlines the indicators returned by the proposed model. Naturally, they focus on mobility and accessibility-related measures, while ecological issues such as the micro-climate, environmental risk assessment, and energy efficiency are not included yet. For these aspects, specialized models exist (also with implementations for Grasshopper). The integration of those as additional modules in the future would undoubtedly augment the toolbox.
The indicators allow making statements on travel time-related figures such as *avg. commuting time to jobs*, *avg. travel times to closest facilities*, *avg. travel times to all facilities*. Moreover, network related dimensions are covered, such as *footfall to facilities* along with *closeness* and *betweenness centralities*. Finally, several indicators deal with points of interests or locations individually, allowing to analyze *visitor numbers, visitor origin, catchment areas*, and *utilization* (which can be interpreted as an implementation of an extended huff model). These indicators alone allow for wide ranges of analysis from location choice optimization of private or public entities to decisions making on transit related issues. More details and examples regarding the indicator system are provided throughout the next chapter.

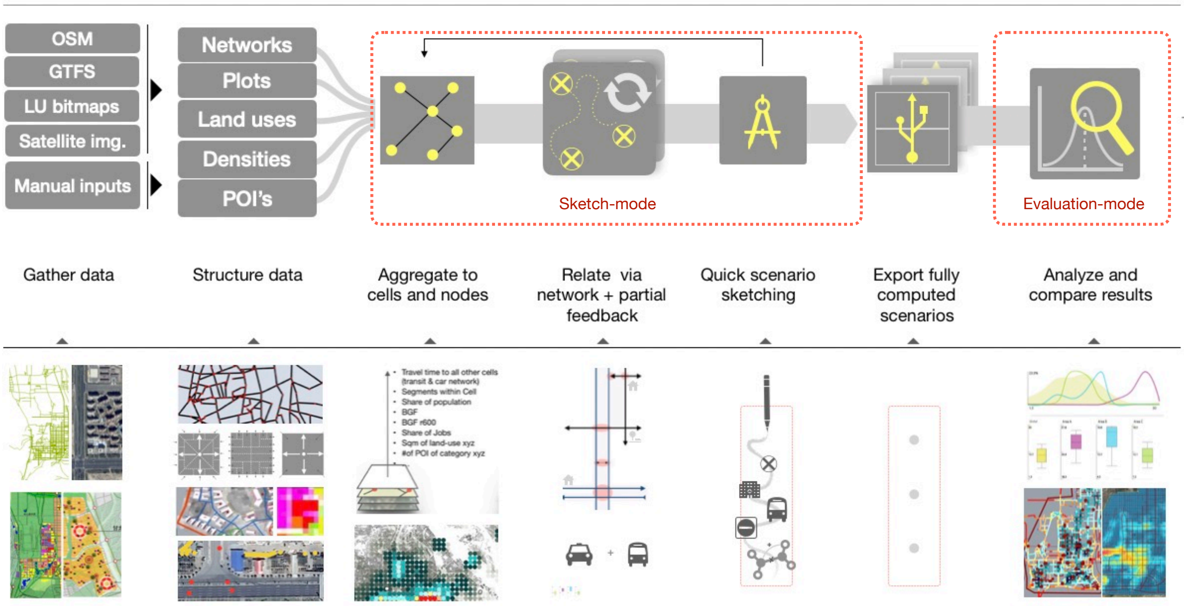
Arguably, each project and use-case demands individuals evaluation criteria. The modular structure and implementation of the toolbox in Grasshopper allow augmenting the model by additional indicators quickly. An example is given for the case-study in chapter four, where more than 15 additional indicators are included.

Spatial unit	Indicator	Unit	Calculation	Example for application
POI	Number of visitors to POI x *	People	As calculated by the gravity function	Can be used to see how visitors all distributed among the POIs of the same category, which might help to find the right location for adding another one to lower use-pressure etc.
POI	Use-pressure on POI x *	sqm per person	Square meters of POI x / number of visitors	This indicator could be used to see how much the addition of a new park in the city reduces the use-pressure on the other parks.
POI	avg. commuting time of visitors to POI x *	Min	Sum of travel time of visitors to POI x / number of visitors	The average travel time to of visitors to the amenity.
Network	Betweenness centrality (several radii)	Score 0-1	Counts how many times a street was passed when taking the shortest-path from each street to all other streets in the network.	Streets with a high betweenness centrality are likely to have a high frequency of pedestrians/cars crossing. Thus, it can be interpreted as a proxy for economic potential for commercial uses or to roughly estimate traffic flows.
Network	Closeness centrality (se-veral radii)	Score 0-1	Computes the total distance (time) from one street segment to all other streets in the network	A high value indicates that many other places can be reached within a certain time. Therefore a high closeness value correlates with the location of centers in towns and can also be thought of as a proxy for economic or retail potential.
Network	Footfall to POI's of category x *	People	Via gravity function and tracing of shortest paths using population in origins as weight	Useful to see which streets are used most often of people on the way to an amenity. This helps to see how new infrastructures affect movement flows and potentially helps to decide on where co-projects such as street renovation should be prioritized.
Analysis grid	travel-time to closest POI's of category x	Min	via shortest-path algorithm	Useful when POI are similar and it doesn't matter which one of them can be reached.
Analysis grid	expected origins of visitors to POI x *	People	The whole population is distributed among all POI of category x based on the POI weight and distance	Visualize where visitors to a POI originate from. Might help to decide on which areas are in greatest need of additional amenities or to balance the visitors among similar POI.
Analysis grid	avg. Commuting time of residents of an area to POI of category x *	Min	Based on the origins of visitors to POI x	Visualizes in which area avg. travel times are lowest and thus in greatest need of improvement measures.
Analysis grid	Total gravity (access/ opportunity) for an area to POI's of category x *	Score	Sum of gravity values of category x POI's in an area	Usually, the more gravity the better, as a high value means more POI can be reached in a shorter time. Helps to see which neighborhoods are undersupplied by amenities or have less overall access to jobs.
Analysis grid	Accessibility to Jobs/Land use/POI	Score	The average of the opportunity in area j (eg. jobs, m2 of land use or POI) divided by travel time between i and j	A measure often applied to estimate the impact of infrastructure investments such as of public transport. Compared to total gravity, the discount on travel time is linear here and the results easier to interpret.
Analysis grid	Avg. commuting time to work of residents in cell i *	Min	Based on the origins(residents)/ destination(job) distribution (via gravity function) and shortest path calculations	An important measure for the overall performance of a city. Note, that an improvement in the network (eg. new tram-line) does not necessarily lower the avg. Commuting time as residents have now more jobs in reachable distance. This indicator should be looked at together with total gravity and accessibility changes (see above).
Analysis grid	Tot. Commuting time caused by jobs in cell i	Hrs	Sum of commuting time from all cells to jobs in cell i * number of people of a cell working in cell i	Areas that cause high commuting times might be looked at with priority when planning infrastructure or traffic improvements. Especially when looking at the origin of commutes at the same time.
Analysis grid	Tot. commuting time spend by population in cell i	Hrs	Sum of commuting time from cell i to jobs in all cells, weighted with number of people.	Areas that have high commuting times might be looked at with priority when planning infrastructure traffic or improvements. Especially when looking at the origin of commutes at the same time.
Analysis grid	Catchment area of POI x	sqm	Via travel time to clst. POI of category x	Provides an overview of catchment areas of each POI

Table 3: Overview of basic indicators returned by the toolbox. Source: Author

3.3 Implementation

The models structure and general workflow is illustrated below (Fig. 4). Each part is described in detail below. Naturally, the first step is feeding the model with data either from external sources or by manual addition/drawing. Next, the data needs to be structured and prepared. In the *sketch-mode* both the models numerical and spatial parameters and configu-



Source: Author

Fig. 4: Model structure and workflow

ration. During this process, the users get constant feedback on selected indicators in near realtime. When a scenario is drafted, it can be exported (with all indicators being computed) as a .ghdata file and, together with other scenarios analyzed and compared in the evaluation-mode.

This section follows the structure depicted above by first elaborating on data inputs, their structuring, and more technical details on the model before an introduction to the user interface is given. Finally, both the sketching and evaluation modes are described in detail, while two brief use-cases illustrate their application in practice.

3.3.1 Data structure, levels of aggregation and scalability

The toolbox operates on several different spatial units and supports interpolating, aggregating and dis-aggregating between them. Most importantly there are network segments, points (for POI and network nodes), shapes (e.g. blocks or plots) and two grids: one high-resolution *analysis grid (20 - 100 meters),* that most other data is mapped on before exporting, and one low-resolution grid (125 - 1000 meters) that is used to compute a distance matrix[7] and enables the model to work with larger cities and areas. Fig. 5 and 6 provide overviews on the interpolation methods between the spatial units used in the model.

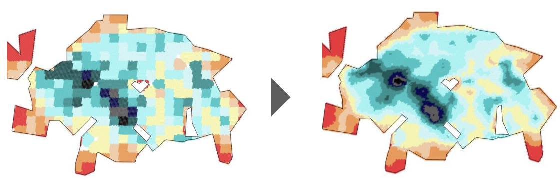


Fig. 5: Bilinear interpolation between two grid resolutions. Source: Author

[7] Low-res grid and distance matrix are exchangeable terms in this section

Different levels of detail can be used at the same time. In cases where all areas have (significant) relation to all other areas, it can make sense to use the reduced complexity offered by adopting the aggregation level of the distance matrix: relevant values are aggregated to the low-res grid cells and then related via the pre-calculated distance matrix. For instance, in order to compute the average commuting time to jobs, all spatial units that have either a population or job number higher than zero need to be included in the calculation. In the case of Vienna (ca. 25000 network nodes), using a grid of 500m (ca. 1200 grid cells) would decrease computing time by up to 500 times. In the case of Kagan[8] (medium-sized city, with 3500 network nodes) the low-res grid was used to calculate commuting patterns to jobs, while network nodes were used when computing travel patterns to point of interests. This functionality of combining different scales and level of details ensures scalability and flexibility of the toolbox.

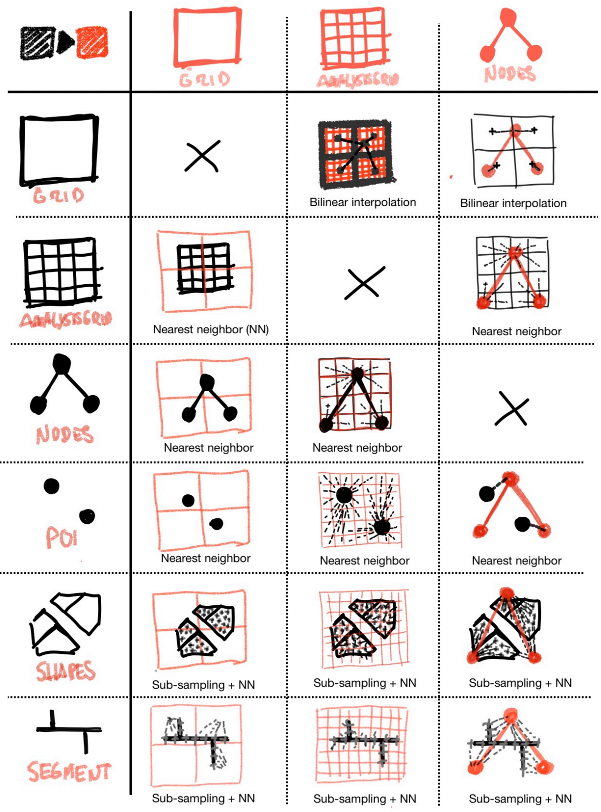


Fig. 6: Interpolating between aggregation levels and spatial units. Illustration of how the model aggregates, dis-aggregates and interpolates indicator values between different spatial units. Source: Author

In order to achieve a better compromise between accuracy and calculation speed, cells are only generated near network nodes and within the defined area of analysis. Furthermore, a script was developed that adjusts the resolution of the low-res grid (or distance-matrix) adaptively. The algorithm compares the travel times from all contained network nodes (in a cell) to a set of test points (e.g., 500 randomly distributed nodes in the network), if the travel times show significant discrepancy the cell gets split into four, and the algorithm starts over. Either a threshold can be defined manually, or a maximum number of cells can be set. For instance, in case a cell covers two sides of a river, the algorithm would divide the cell until no nodes of the other side are within one cell, which arguably increases accuracy by a good deal (see Fig. 7).

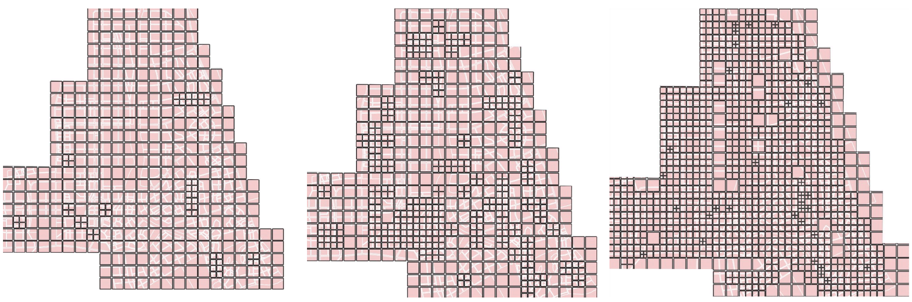


Fig. 7: A generated low-res grid. The initial cell size is ca. 1000 meters. From left to right, the threshold is increased, thus more and more cells are divided once (x4) or even twice (x16). Source: Author

[8] Parts of the toolbox were tested and developed along with a project in the city of K. Thus, below the town will be repellently used as an example.

3.3.2 Data Inputs

Management of data inputs

All input data and design related user inputs are currently managed and structured via the layer system of Rhino. To a degree, the system is dynamic and customizable, as the user can add sub-layers to sections such as Land use, Point of interests, roads, etc. and thereby create new classes for which individual parameters can be defined (such as avg. speed, gravity model parameters, etc.). This method can also be used to test different parameter settings within a single model run. Fig. 8 gives a detailed overview of the layer system and supported data types.

Vector data

Transportation network: Vector data can be supplied either manually, directly in the Rhino canvas, or by importing them as Shapefiles. For OpenStreetMap data, the user can define which road categories to use and how to remap them to match with the structure of the model (e.g., in case less or different road categories are planned to be used later on). In the layers, *manual_roads*, users can sketch their networks, fill out missing areas, or propose new designs and infrastructure elements such as highways or bridges. Each sub-layer of *manual_roads* defines the class of the curve added to the system, individual parameters such as average speed, waiting, and exit times can be defined.

To model transit, simple geometries that resemble the transit routes can be imported e.g., from shapefiles. Ideally, one transit line is represented by one curve. For more technical information, see section 3.3.3.

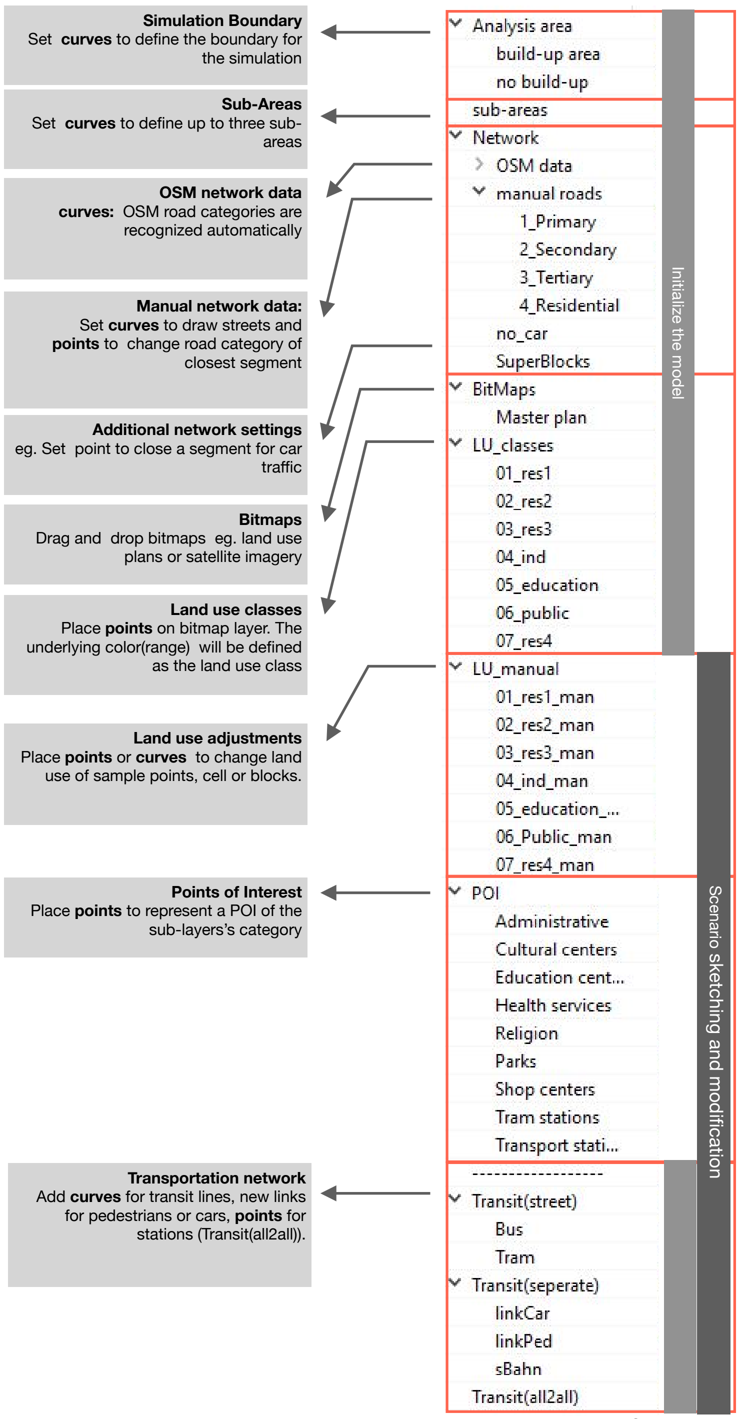


Fig. 8: Layer structure and user inputs

Source: Author

Points of interest: Similarly, points of interests need to be structured into classes or categories. The gravity function is applied to each class individually. Thus parameters need to be defined for each class (Fig. 9).

Blocks and plots: Blocks are generated based on the pedestrian street network using a component from DecodingSpaces. They are most importantly used during sampling land use data from bitmaps, as will be described below. Moreover, the square meter of a block can optionally be used as default weights for POI's located in them.

Bitmaps

A part of the Grasshopper definition was set-up to deal with defining land use classes from bitmaps based on color ranges. For each class, a sub-layer in *LU_classes* is created and named. Next, the user needs to add a bitmap on the layer *BitMaps* (for instance, a land use plan). Classes are defined by adding a point or points to the sub-layers of *LU_classes*. For example, to build a class for educational land uses, a point needs to be placed on the bitmap with the corresponding land use and color code on the layer *LU_classes::education*.

There are several mechanisms to cope with inaccuracies and color variance emerging from compression or other reasons in order to increase the robustness of the classification process. By placing multiple points, the user can define a color range instead of a single color that is searched for. Additionally, a threshold or tolerance can be defined. This allows classifying sample points even if they do not perfectly fit into the color range.

Finally, the bitmap is sampled, relying on a point cloud generated within closed curves (see Fig. 11). The most common use-case is to use blocks or plots for the sampling of a land use plan. Each point gets tested for whether or not it can be matched to a class. This is done by applying the closest point method (The RGB colorspace is used, which can directly be translated into x, y, z coordinates) and testing for thresholds (see above). For each curve, it is tested how many points of which land use is within its boundary. When a land use class surpasses a threshold (say 5%), it is included when calculating the total land use shares for the block/plot. Therefore, even if the original land use plan has overlaying symbols and texts, blocks and parcels can rather accurately be classified.

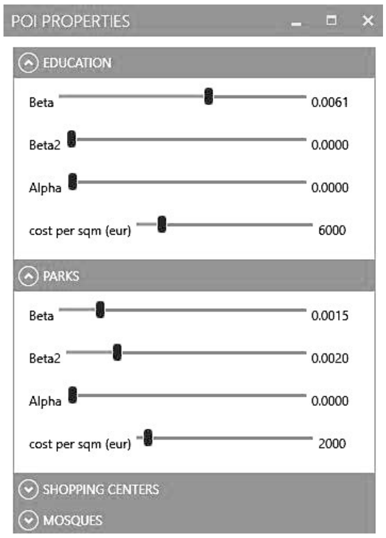


Fig. 9: POI category parameters. Source: Author

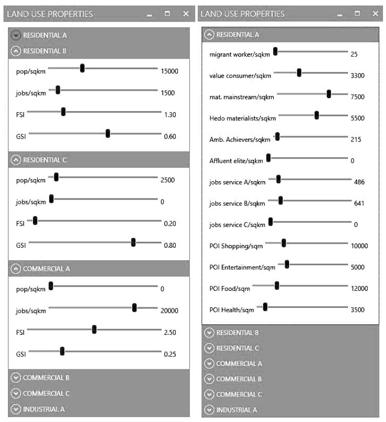


Fig. 10: Land use parameters. Left: set-up used for the case-study; right: hypothetical, more complex set-up. Source: Author

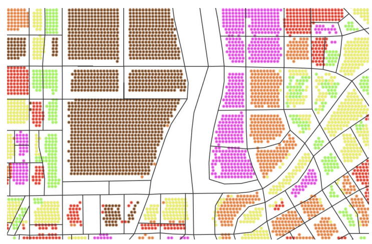


Fig. 11: Sampled points in blocks Source: Author

For each land use class, several parameters need to be defined. Most importantly, the number of jobs and population per square meter. However, more complex set-ups are possible as well as demonstrated in Fig. 10. For instance, with corresponding data available, it can be defined in which typology which social milieus, income groups or types of jobs (e.g., service, industry) are represented in which share. This can later be employed for user-group specific impact analysis (using the spatial filters, see next section) and planning.

Georeferencing and projection management

In case geo-referenced data is used Plug-ins such as DecodingSpaces provide components to manage projections and transformations.

3.3.3 Network and graph modeling

All data provided in sub-layers of Network and Transit are used to create a transportation network graph. Both, one for pedestrians/transit and one for cars. Junctions act as nodes and segments as edges. Each edge and node have a set of additional information associated with. For instance, the edge category (e.g. pedestrian, Metro line 3, Tram line 12) depends on the layers (or OSM attributes), geometries were placed in. These attributes can be used to define the cost for edges. This is generally done by dividing their length with the avg. travel speed associated with their category (Fig. 12).

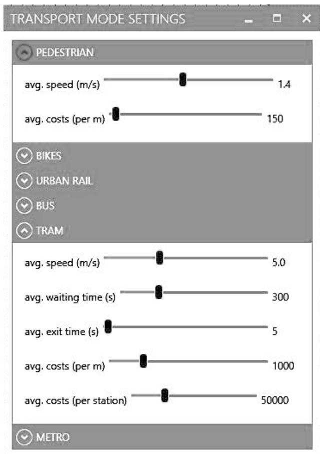


Fig. 12: Transport mode and road category parameters. For each category avg. speeds and waiting times can be defined. Source: Author

Before the graph is created, the network geometry needs to be processed, cleaned up, and simplified. The user can adjust several parameters to alter the degree of simplification (see Fig. 13).

After the original network geometry is simplified and broken down to segments, a graph for the pedestrian (or car) network is generated, using the plugin *Spiderweb*[9] which returns a basic, directed graph as a data tree. This is the starting point for customizing and merging graphs generated for transit with them. The chosen approach is to generate graphs for the transit network (one for each transit line) individually. In the next step, the transit graphs are merged into the pedestrian network (see Fig. 14). This provides a convenient way to include average waiting, exit, and transfer times for different transportation modes while keeping the graph compact and simple.

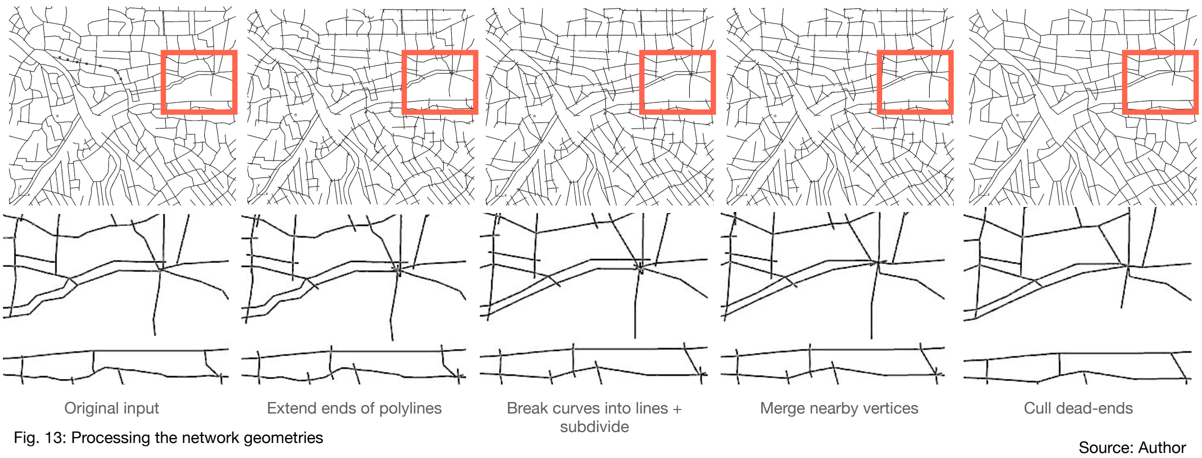


Fig. 13: Processing the network geometries

Source: Author

[9] *Spiderweb* is a plugin for *Grasshopper* developed by Richard Schaffranek for graph processing and analysis.

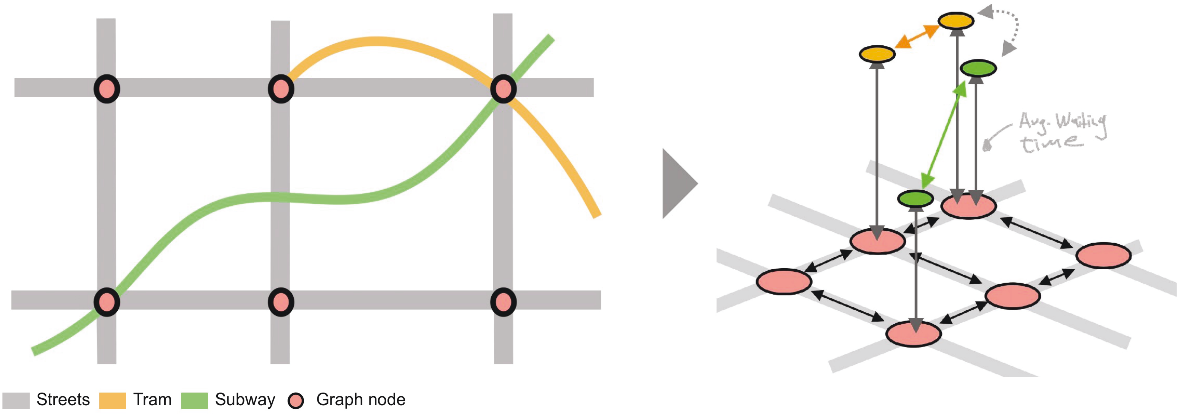


Fig. 14: Graph structure of the merged pedestrian and transit network
Left: Urban network configuration including roads a tram and a metro line. Right: Adjusted graph structure which allows to define average wait and exit times for each transit mode. Source: Author

The car network is modeled very similar to the pedestrian network - naturally with different avg. speeds for the segments. Different kinds of roads (and speeds) can be defined through the sub-layer system or, in case open street map data was used, using the attributes assigned to segments (Fig. 15).

In some cases, it can make sense to model the car network in greater detail. Fig. 16 illustrates an example: First of all the basic graph returned by spiderweb is used and then altered with a Python script that, according to a set of rules, changes the connection possibilities. In the example, a total of six different junction categories were defined, such as a T-junction between a regular and one-way road, u-turn possibilities to mention a few.

Sometimes more options for the graph creation are useful for case-specific fine tuning, such as closing road segments for cars or changing their category afterward either to fix data errors or to test scenarios. This can be done placing points (instead of curve geometries) in the sub-layers of the network

Fig. 15: Various road categories in Vienna.
Colors represent a road category. Light blue represents residential roads, dark violet roads of highest hierarchy. For each type, individual properties such as avg. speed can be defined. Source: Author

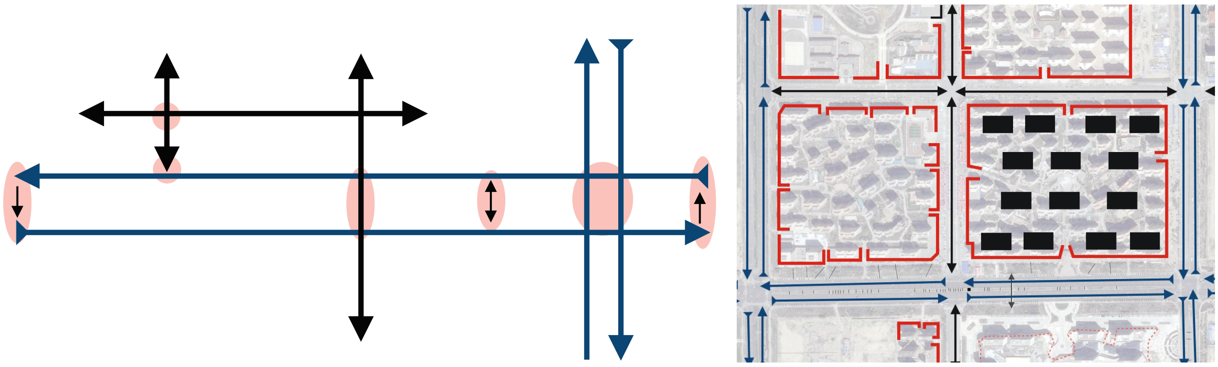


Fig. 16: Adding details to the network graph: Left: Six intersection types & connection rules to model one-way roads, u-turn possibilities etc.. Right: Structure of the city, the six intersection types were derived from. Source: Author

Network analysis and shortest path calculations

After the network graph(s) is set-up, it can be used for analysis, routing, and travel time estimations - it represents the foundation for most of the evaluation indicators the toolbox computes. In most cases (visitor distribution and footfall to POI, travel time radii for dynamic areas, etc.), components of SpiderWeb in combination with custom Python scripts are used. The population distribution is usually used as a weighting factor.

To compute closeness and betweenness centrality for small and medium cities, components of the DecodingSpaces plug-in are employed. However, for more extensive networks, the system tends to run out of memory or crashes. Thus, alternatively, a stochastic approach was implemented using SpiderWeb components for the shortest path calculation in combination with Python scripts for random sampling and path tracing. Tests showed that even relatively low sampling rates (randomly distributed pairs of origins and destinations) performed reasonably compared to precise calculations. Fig. 17 shows the results of a test run for the city of Weimar (ca. 950 segments). By just computing 5% of all possible origin/destination pairs, an r-square of 0.76 was estimated, with a computing time of 0.4 seconds (a speedup of 100 times). A sample size of 20% took 2.5 seconds corresponding to a 16-times speed up and with an r-square of 0.95. Instead of randomly, samples could also be generated based on weights (e.g., population and job densities). The test suggests that estimating betweenness delivers reasonable results concerning the savings in calculation time - at least for initial, explorative design phases and when quick feedback is necessary. The user can set a sample rate for the calculation of footfall to the point of interest (by default, it is set to 90%).

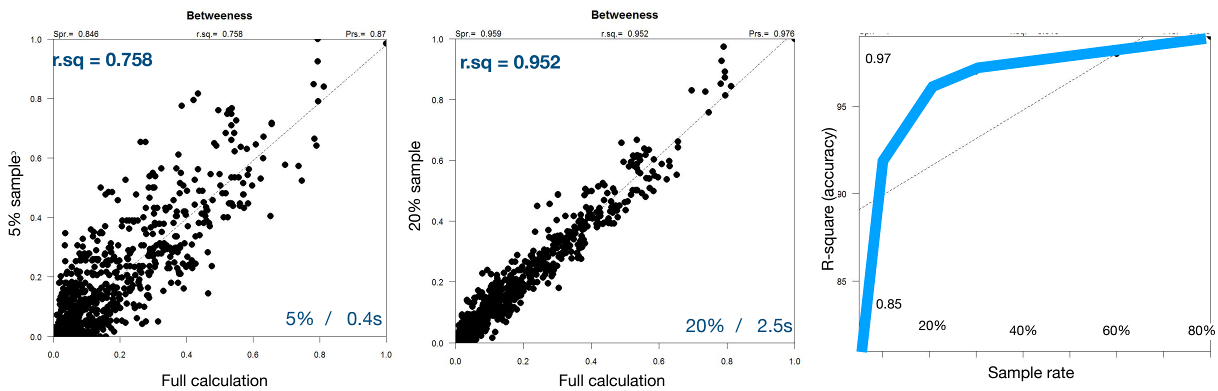


Fig. 17: Estimating betweenness. From left to right: 1 and 2 show scatter-plots with accurately computed betweenness values on the x-axis and an estimated betweenness using a 5% (1) and 20% (2) sample. 3 depicts the relation between the sample rate (as a fraction of the number of samples required for the accurate calculation) (x-axis) and the accuracy or r-square. It suggests that after a sample rate of 20% additional gains in accuracy come with a relatively high computational costs. Source: Author

Short comings and future improvements

The presented approach to model transit has been tested using the city of Vienna. The model's results were compared with travel times computed via google maps with overall good results (see Fig. 18). The most considerable differences/errors between the model presented here and google maps were found in road-bound transit options such as bus and tram in remote areas, with arguably less traffic and broader roads. This could be roughly equalized by adding a factor affecting the average speed depending on e.g., the local built-up density around the segment, its distance to the city center or centrality values, etc., while at the same time keeping the average speed steady. A genetic optimization algorithm such as

Grasshoppers Galapagos could be employed to estimate better weight for such a parameter by trying to get as close as possible to the travel times of a series of sample points computed by google maps.
Additionally, some tests were conducted by building a simple, stochastic traffic model for car and transit networks. The distribution of trips is handled probabilistically, employing a gravity function. Different destination types can be defined and modeled. Fig. 19 gives a very brief overview. Despite the low level of detail, the calculation time for Vienna was quite high (ca. five hours), which is why this module is not an integrated part of the toolbox so far. However, especially in cities with sound transit networks, having (a rough) traffic model like this is crucial to improve estimating pedestrian flows or their potential.

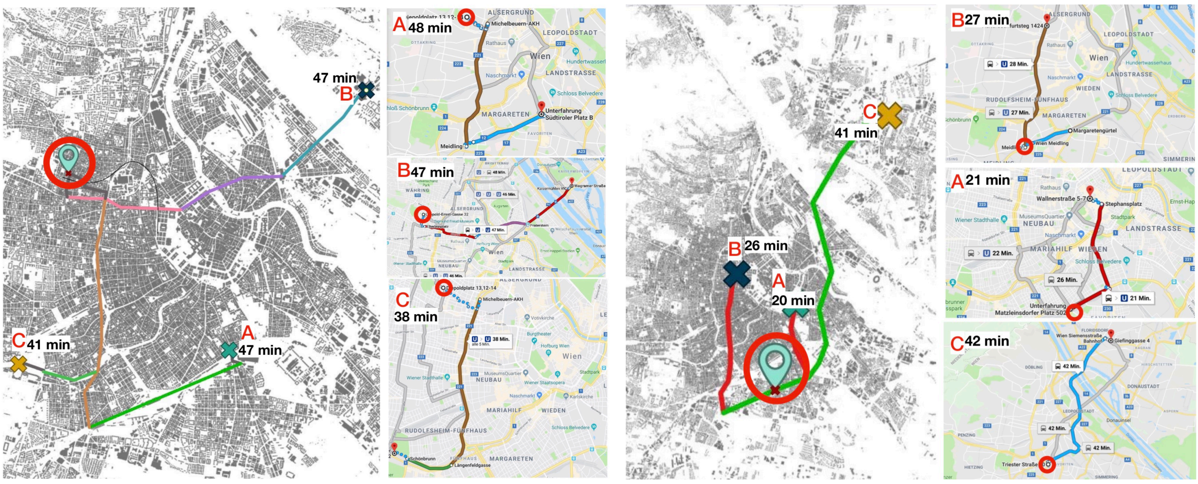


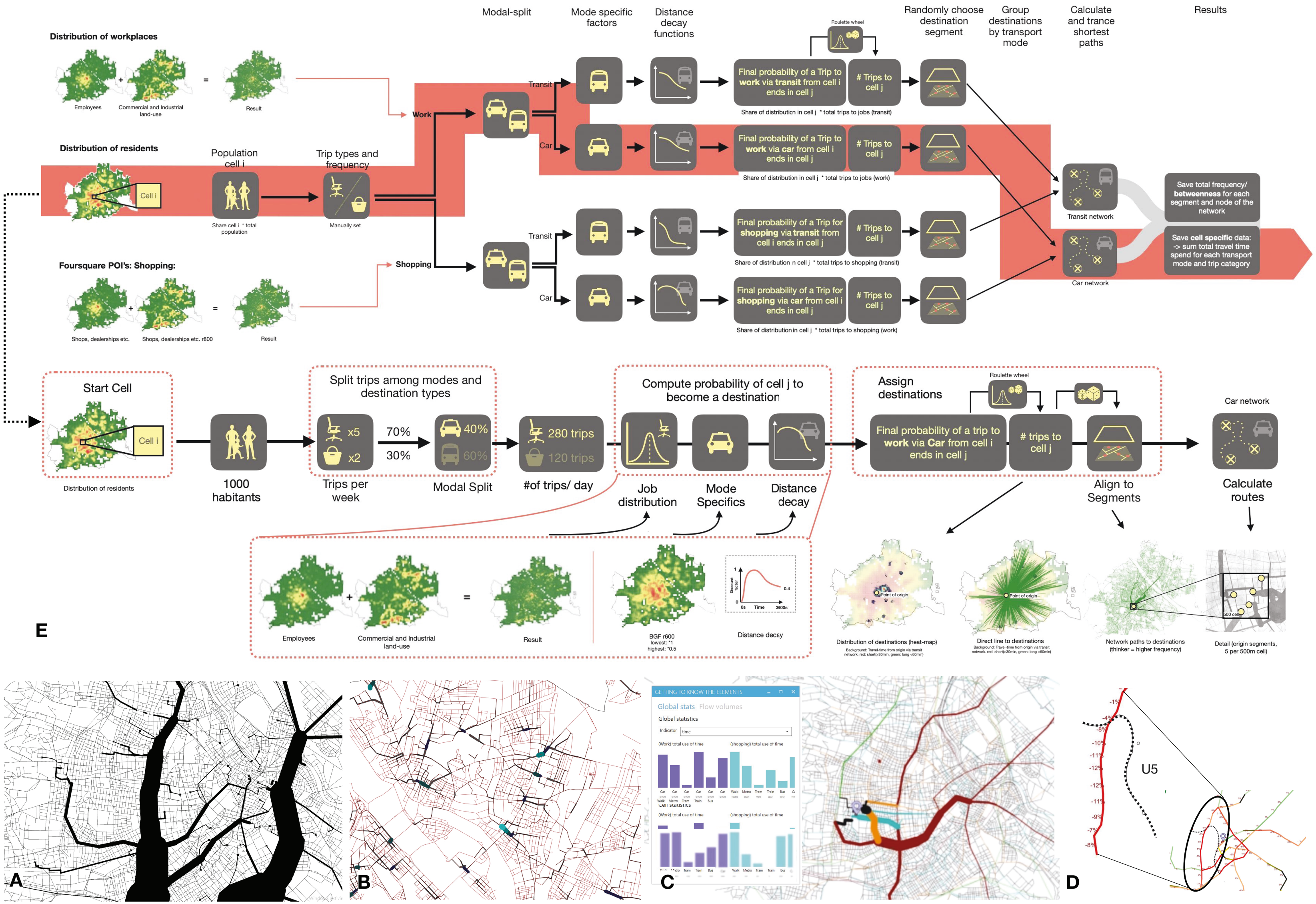
Fig. 18: Comparing travel times calculated by the model with google maps. Source: Author

3.3.4 Gravity model implementation

The gravity model is implemented as described in chapter two, formula 5. The function is used to 1) decide whether trips from an origin to a destination are made at all and 2) if made, to which destination they are routed. In short, the function distributes trips between origins and potential destinations in a probabilistic fashion.
The gravity function works both with points of interest and data aggregated to the low-res grid (distance matrix). For instance, in the case of Kagan, calculations for commuting to jobs were conducted using the low-res grid and an underlying distance matrix.
For each POI category, parameters of the gravity function can be defined individually. If no data is supplied the values are set as follows: beta = 0.0022; beta2 = 0; alpha = 1. The default value for beta is commonly used to model pedestrian movements (Bielik et al. 2018). In case beta2 equals zero, the term will be disabled for the calculation, which means that the whole population is distributed among the destination classes.
For origins, the population number is used for weighting by default. For destinations, the size of the block a POI is located in is set as the default weight. However, this can be altered manually. In case the low-res grid is used, weights are derived by the square meter of land use, the number of jobs, population, etc. - or other values that had been defined during the setup. This also allows conducting user-group specific analysis, if population numbers have been split up into groups during the setup process.

Fig. 19: Concept for a static traffic model. The traffic modelâ€™s structure is illustrated below (E). Trips are distributed probabilistically and stochastically as a function of the distribution of retail, jobs, and the distance to them (from points of origins, weighted with the population). The model is static. Effects such as congestion are not yet considered. Moreover, passengers or cars do not affect each other as the model is not agent-based. To streamline computation speed in Grasshopper, the model works on a resolution of 500-meter cells, and sub-sampling to five randomly selected segments to map trips between specific network segments.

- A: The segment width represents passenger volumes. Apparently, the subway and light rail handle most of the passenger streams.Â
- B: Transit segments are hidden for visualization: pedestrian movement patterns, shaped by the transit network, emerge.
- C: Evaluation of trips from one 500 meter grid cell to jobs. Segment colors represent transportation mode. Bar plots show how much time was spent commuting per transport mode and trip type (work and shopping).
- D: Estimated difference in ridership levels per segment of Subway line six after adding a new subway line (U5).



Source: Author

The gravity function is employed to calculate various indicators such as footfall and visitors to POI, avg. commuting time and more. Thus, basically representing an extended version of the huff model that can be used for, but is not limited to, commercial location choice problems.

3.3.5 Distribution of jobs and population

Two of the most essential variables in the model are the distributions of population and jobs. Thus an approach to determine them is presented in this section. Often there is no data available with sufficient spatial resolution. Several approaches exist to estimate these critical metrics with proxy data. One approach that is especially useful for this work is described in Jayasinghe et al. (2017). They used the street network and centrality measures to estimate the distribution of jobs. Logically many firms move to cities because they rely on access to workforce and talent. Thus they can be expected to make location choices in areas with high closeness centrality. Additionally, firms that are dependent on footfall and benefit from high pass-by frequencies can be expected to prefer locations near places with a high betweenness centrality (see Yang 2015). However, what is not captured by these measures are companies acting upon different criteria for location choice as common for industrial entities. Therefore it is beneficial to combine the described approach with land use maps (if available). Fig. 20 provides one example of how different measures can be merged to conclude to a job distribution. Weights for each input could be estimated based on the distribution of jobs between sectors such as industrial and service.

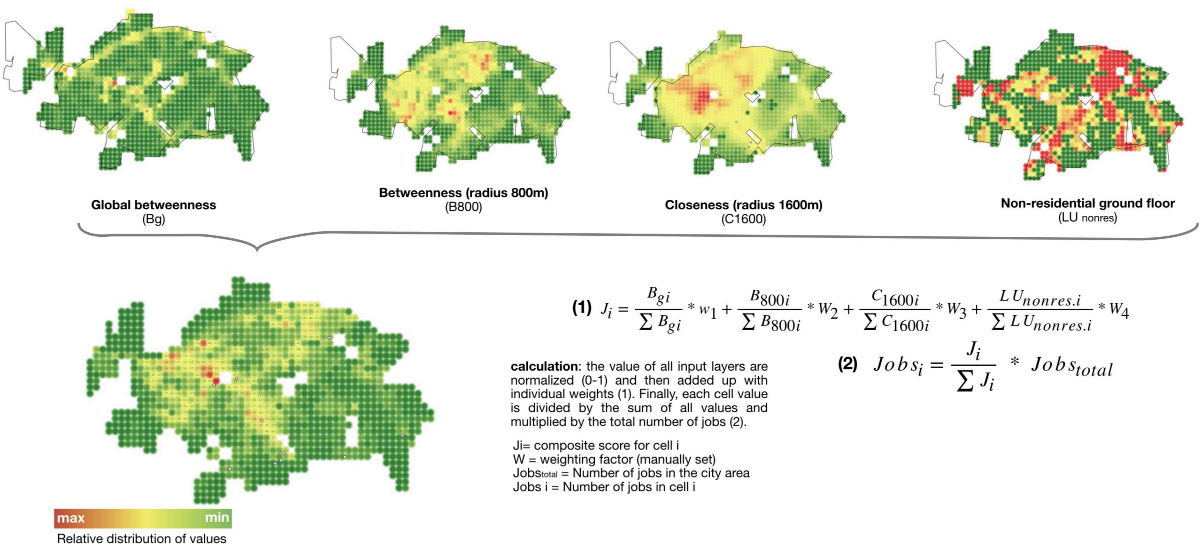


Fig. 20: Approach for estimating job distribution

Source: Author

Estimating the distribution of the population is usually easier. In case a land use map is used, each land use class needs to get assigned with a number on its avg. population density. In case building geometries are available (either footprints or gross floor area), those can be used alternatively (see Fig. 21 for more details).

Even though this approach rather resembles an "educated guess" than delivering accurate numbers, one can expect that even these rough estimations improve the accuracy of the model.

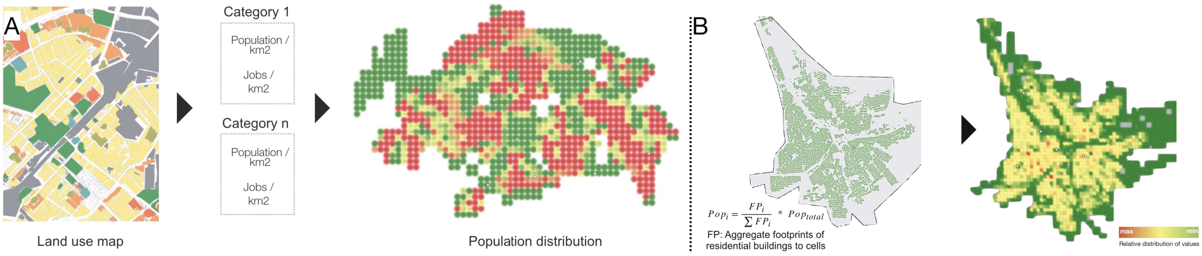
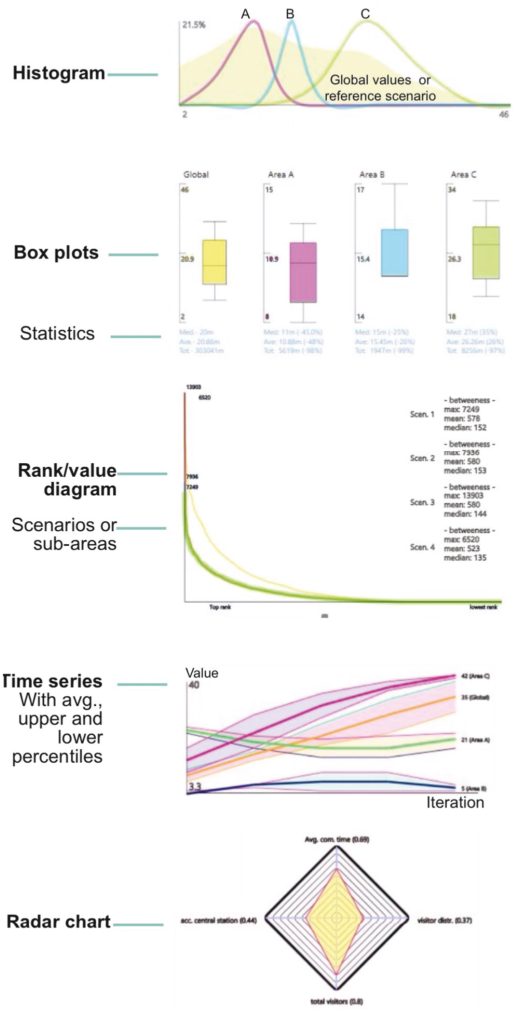


Fig. 21: Approaches for estimating population distributions. A: Method is based on land use maps. For each land use, category population densities have to be defined. These numbers are then used for the estimation. B: footprints or gross floor area of non-residential buildings (in this case, data was supplied by the project partner) are aggregated to spatial units. Finally, the relative share per spatial unit is multiplied by the overall population. Source: Author

3.4 Application: User interface and functions

As data inputs and scenario sketching are mainly managed via the layer system of Rhino, the graphical user interface focuses on navigating through the indicators. The sketch- and the evaluation-mode have separate user-interfaces that share a similar logic and are made-up from the same set of components for mapping and chart plotting. These components can be combined and rearranged to expand the system flexibly by e.g. new indicator categories. Fig. 22 provides a brief overview.



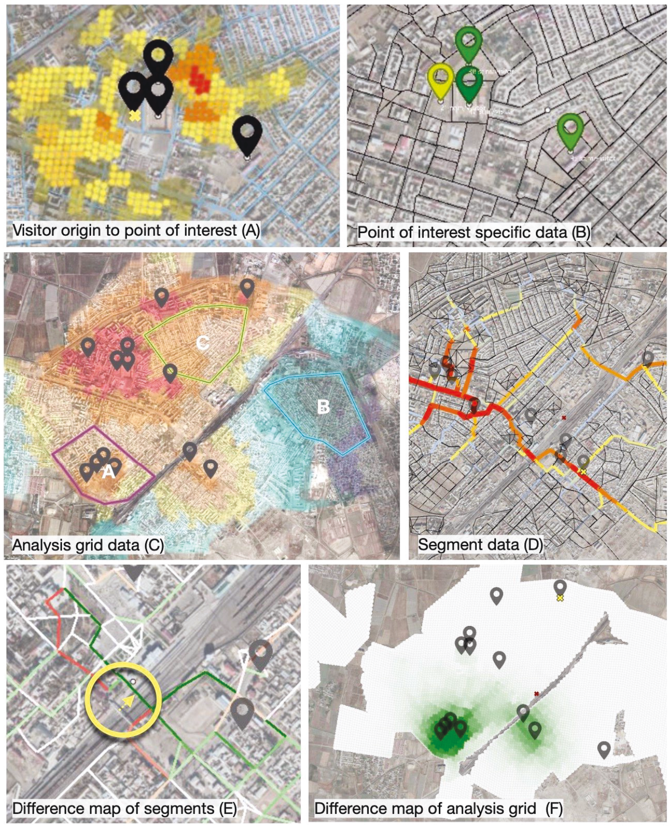


Fig. 22: Overview of basic diagrams and mapping functionality.
Left: Overview of diagrams implement for the GUI of the toolbox. Right: overview of the basic mapping functionality of the toolbox.
E: shows difference-map of footfall to major shopping localities after changing the position of a pedestrian railway overpass.
F: The difference-map shows the change in accessible square meters of green space discounted by with a distance decay function after the addition of six parks. Screenshots from toolbox. Source: Author.

Overall the system is streamlined to support comparative analysis: between the overall situation and sub-areas in the sketch-mode, and between scenarios in the evaluation-mode. Usually, for each indicator, the values for either all areas or scenarios are displayed at the same time. Moreover, the relative difference to a selected reference scenario or area is displayed in percent. In terms of mapping, difference-maps can optionally be shown (see Fig. 23).

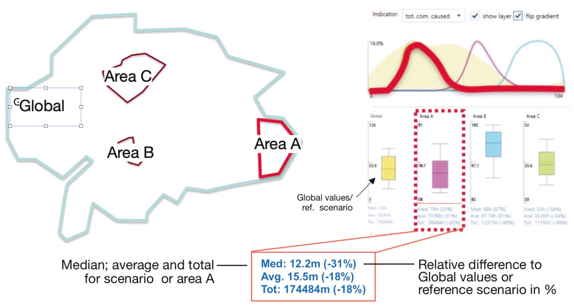
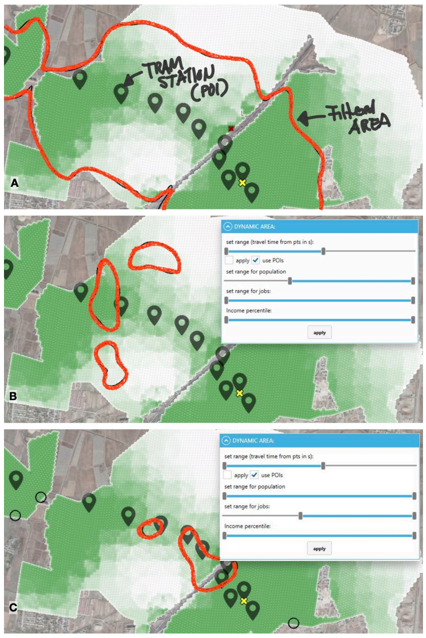


Fig. 23: Interface for comparative analysis. Diagrams and indicator are arranged to streamline comparative analysis between selected or filtered sub-areas (sketch-mode) or scenarios (evaluation-mode). Source: Author

3.4.1 Spatial filtering and sub-areas definition

The toolbox provides the functionality to define up to four sub-areas to analyze the impact of a proposed intervention on a local level. These areas can either be characterized spatially (by sketching a curve around the area of interest on the layer *sub-areas* or by defining dynamic areas (Fig. 24).

Dynamic areas are rule-based selections. Only spatial units (such as nodes or cells) that meet all of the selection criteria are included. Such as the population, jobs or income, social-milieu, etc. (if data is available). This allows to keep track of, let's say, how low-income groups benefit or are disadvantaged by an intervention in particular. Below are two examples:



Different settings for the dynamic area definition. The filtered area is represented by the red curves. Screenshots from the toolbox.

A: *use POI* (as origins) is activated, the filter is set to sample all data points within a 15m walk.

B: another condition was added, that is only areas with a high population density.

C: The condition for a high population density was replaced with a minimum requirement toward the number working places.

Source: Author.

Fig. 24: Settings for dynamic areas

Fig. 25 left: In this example, the dynamic area filter is set to select all places within 15 min. travel time from the railway station. The filter is applied to all four scenarios (evaluation-mode) individually. A curve outlines both the selection for the reference- and the selected/displayed scenario. In this case, the actual situation is compared to a scenario with a tram line. Right: three manually drawn polylines define three sub-areas. Histograms and box plots show the data distribution for the global as well as local context (sketch-mode). Additionally, the relative difference of key statistics to global values is shown.

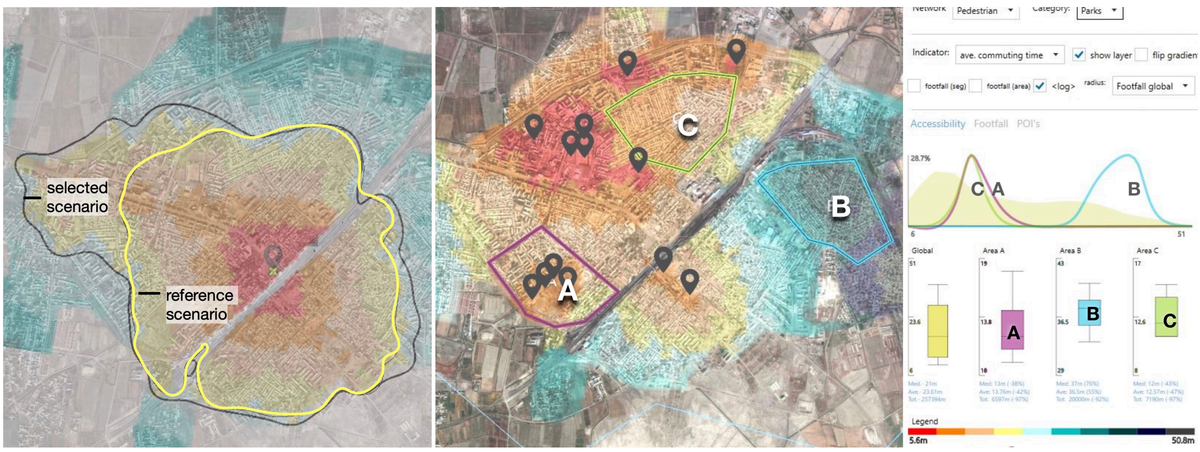


Fig. 25: Sub-area definition. Left: Dynamic area filters are set to return all areas that are within a 15 min walk from the railway station. The filter is calculated for all four scenarios individually (evaluation-mode). Right: Three sub-areas were defined by curves. The charts on the left show the values of these areas in comparison with global values (sketch-mode). Source: Author

Fig. 26 exemplifies how looking at sub-areas individually can help to conduct welfare and impact analysis. In this example, the new cultural center in old Kagan (Area C) improves access to similar facilities significantly by a shortened avg. travel time of -33%. The histogram shows that now more than 50% of the population of old Kagan can reach a park/ square in less than 15 min. However, the least connected part (Area B) remains unconnected with no significant improvements. Citywide, the average reduction in travel time is around 23%.
This filtering function offers useful possibilities for more detailed impact analysis and to operationalize user-group sensitive planning in case, socio-economic data is available.

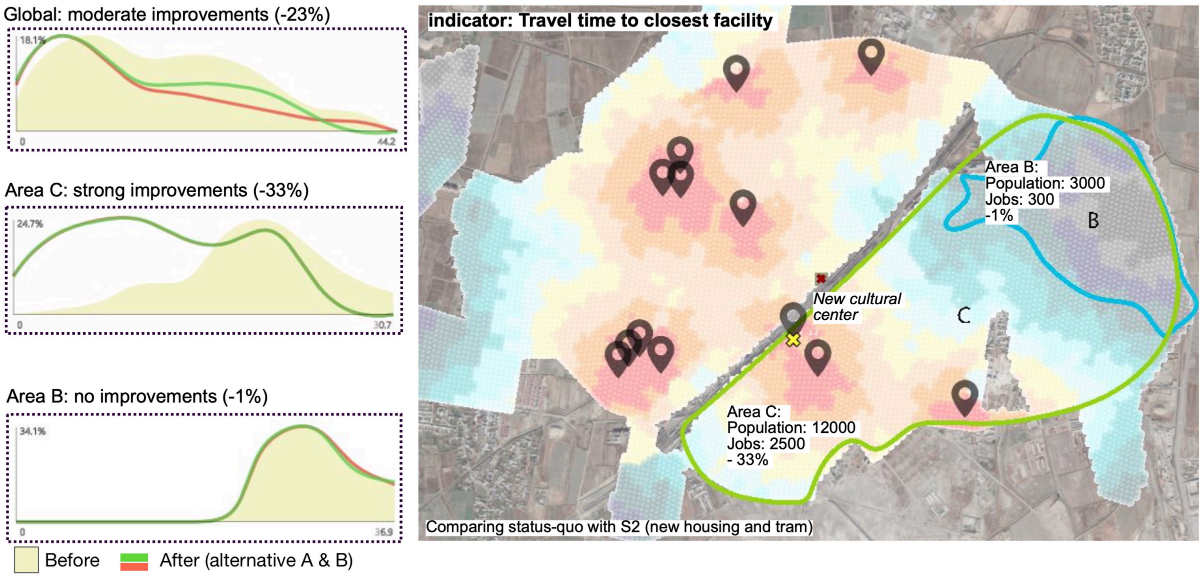


Fig. 26: Applying sub-areas for scenario comparison. Analyzing global and local impacts on access to the amenity type *public spaces* after adding two facilities. Source: Author

3.4.2 Sketch-mode

This mode serves to sketch scenarios and tweak design solutions. As explained, the sketch-mode facilitates preparing multiple scenarios that are then fully computed and ready to be reviewed in detail using the Evaluation-mode (see below).
Here the user can deactivate the computation of indicators and measures that take a long calculation time in order to get instant feedback when trying out different configurations. Thus, facilitating an indicator-guided or evidence-based design process.

For instance, the user wants to fine-tune the location of a new park. In the interface, she selects POI category parks and Indicator avg. commuting time to parks to be displayed and calculated. As different locations are tried out, only this indicator is re-computed to ensure near real-time feedback. Again, values for defined sub-areas or dynamic selections can be displayed at the same time. In cases where the impact for a particular quarter or socio-economic group is of interest this aids in setting the right focus

User-Inputs

Currently, user-inputs are handled within the Rhino viewport and its layer-system (see above). For future iteration of the toolbox, it would potentially be more user-friendly to organize inputs exclusively via the interface. However, a majority of the tool's target group can be expected to be familiar with CAD software and the concept of layers. Therefore, the current implementation should work satisfactorily. The following input channels allow the user to translate their design ideas into the model:

Land uses

Each previously defined land use category has its layer. The user can either place a point or draw shapes/ curves. Points have to be placed within blocks (or plots, cells). They change the land-use of the entire block to the one corresponding to the layer the point is assigned to. Shapes can freely be drawn and overlap with multiple blocks or cover parts of one. For example, if a shape in layer residential B covers 50% of a block with an original land use mix of 20% *commercial* and 80% *residential A*, then its land use mix will be changed to 12.5% *commercial*; 25% *Residential A* and 50% *residential B*. If the shape covers less than a certain threshold (e.g., 5%) of a block, the land use of this block will not be affected.
If previously not existent land uses are planned to be added in a design (for instance high rise residential towers), they anyhow have to be initially set-up in the very beginning in order to not confuse the data structure.

Point of Interests (POI)

Points of interest are structured along with categories represented by layers (e.g. parks, tram stations, schools). It is possible to both add new POIs or remove existing ones. However, the categories used should be defined before setting up the model and be maintained during the sketch-mode.

Setting parameters: For each POI and each POI category, there are specific parameters to set (for each category: beta, beta 2, alpha; and weights for each POI). By default, POI weights equal the square meter of the block or plot they are located in. In some cases, it is useful to manually adjust to the block size in order to represent the actual footprint represented by a POI more realistically. This can be done via the layer *BlockSeperation*. Alternatively, per category, all weights can be set to one or individually manipulated.

Transportation network

There are two ways of changing the transportation network. First, one could make changes in the same way as adjusting the initial set-up of the model (sub-layers of manual roads). However, this will trigger the complete recalculation of a range of algorithms from simplifying the network to re-computing the graph, etc. which can consume considerable time.

To prevent this from happening, an alternative was implemented. Here, only the graph, not the geometry is altered along with interventions. Several types of adjustments are possible:

Changing the road category: place a point near the segment that needs to be changed in a sub-layer of *manual roads*. The segment will change to the sub-layers category. This is useful to test the effect of slowing down traffic on the road or reduce its capacity amongst others

Closing a segment for cars: adding a point near the targeted segment in the layer *no_car.*

Adding new links for pedestrians, cars or new transit lines that run on newly constructed routes independent from the existent transport network (e.g. subway): For either case, a line or polyline is needed as input on the corresponding sub-layer of transit (separate). Stations/connections are created using end-points (and sub-points in case of polylines) with nodes of the existent street network (Fig. 27). This also allows modeling bridges or tunnels.

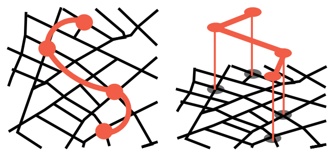


Fig. 27: Adding links to the network .(1). Left: user input. Right: Visualization of graph structure with added subway line (in red: added nodes and edges). Source: Author

Adding new transit lines that run on the road network (Layer: Transit (street)): To add a new transit line (e.g., bus or tram), the user has to draw a polyline on the sub-layer representing the transit mode. It is essential that the start, end, and interim points are placed where stations are meant to be located. In the background, the shortest path between stations is calculated via the street network. The calculated total distance between one-stop and another is then divided by the average speed of the transit mode. Finally, an additional edge directly between the station nodes is added, which is then connected (edge weight represents *avg. waiting time* and *exit time*) to the existent street network graph (Fig. 28). This method allows for both a slim graph structure and quick sketching, as only the stations need to be defined.

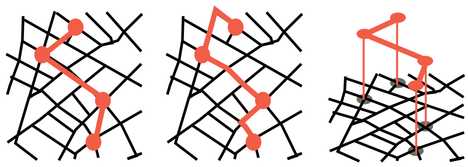


Fig. 28: Adding links to the network (2). Left: user input. Middle: background calculation: shortest path via road network from one station to the next Right: Visualization of graph structure with added bus line (in red: added nodes and edges). Source: Author

Station based sharing: the last mode is suitable to model transportation modes such as station based share-services of bicycles or cars. The user has to place points in the layer *Transit(all2all)* that represent stations. In the background, the shortest path between each station to all others is calculated, stored in a simplified graph and then connected to the main street network (similar as described above) (Fig. 29). Again access and exit times can be modeled (e.g. the average time of unlocking and locking the bike).

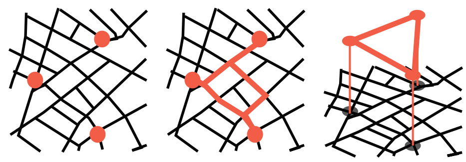


Fig. 29: Adding links to the network (3). Left: User inputs, points representing stations; Middle: background calculation: Shortest paths from all to all stations; Right: Visualization of graph structure with city-bike stations (in red: added nodes and edges). Source: Author

Parameters: Moreover, it is possible to explore the impact of changing parameters such as driving speed for a specific road category or the average waiting time or speed for subways, buses, etc.

By adding sub-layers (and setting the parameters) to a corresponding transit mode, one can add as many different transit types as desired (or define e.g., bus lines with different avg. speeds and frequency).

Feedback and Indicator guided sketching

The added value for the design process stems from immediate feedback given by indicators and maps displaying aggregate values and their distribution - both spatially and statistically. The basic work-flow is structured as follows: First of all, the designer defines indicators important for the (sub-) decision or intervention targeted. Then multiple Indicators are selected, displayed, and tracked within a dashboard. The number of indicators is a trade-off between more information and computing speed.

Interface

The dashboard consists of two sections (Fig 30): On top of the dashboard, a composite score is being shown and tracked over time (in case multiple indicators were chosen). To compute this score, weights and bounds have to be defined for each indicator (this can be done in the set weights tab. Therefore the user retains an overview of how his or her proposal is performing as further developing it. Moreover, for the current state of the draft, a radar-chart is computed that shows how each indicator contributes to the overall performance score individual. A slider allows to browse back to previous states of the design and view its scores in the radar chart.

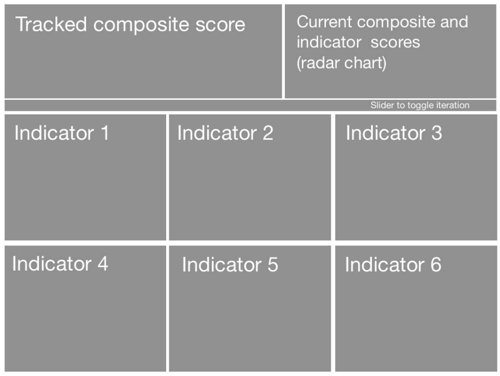


Fig. 30: Interface structure of the sketch-mode. Source: Author

In the lower section, the user can create an individual setup of indicators and their visualization to receive feedback on the current design. Between one and six indicators can be selected to be displayed and computed. A histogram and box-plots are calculated for each indicator and sub-area defined. A text box provides information on the mean, median, and total value of the indicator as well as the population and number of jobs within the areas. The number in brackets shows the change in percent between the current situation and either to the reference scenario or a comparison between the global and area values of the current situation.

Additionally, the mean value of each selected indicator (Global and local, if areas were defined) is being tracked as the user experiments with different options to help to sustain an overview of the different solutions tried. If the mode is switched to *relative* the changes in percent relative to a set reference scenario are computed instead.

Fig. 31: Simplified interface. Source: Author

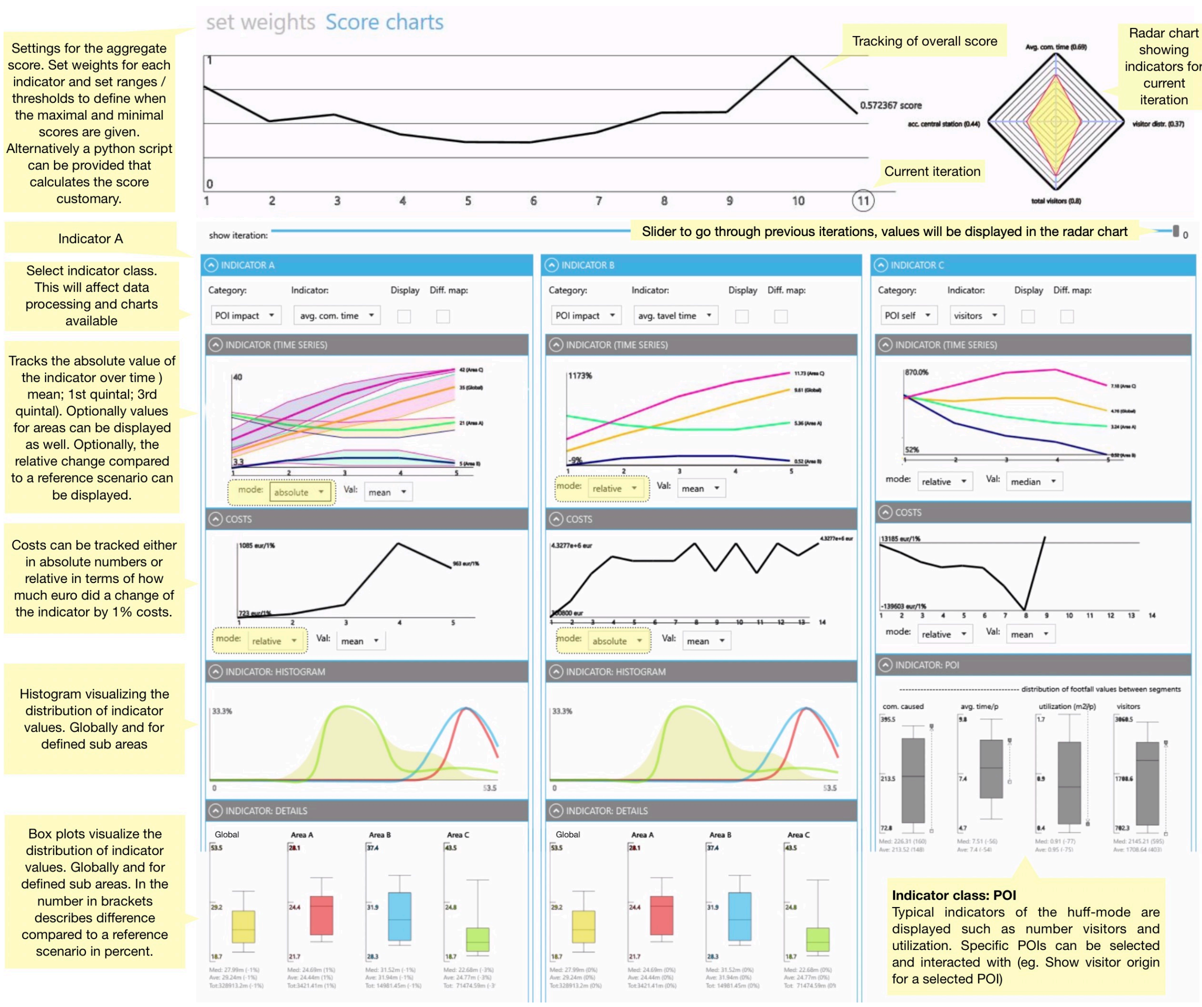


Fig. 32: Sketch-mode indicator monitoring set-up. Exemplary setup for sketch-mode enabling indicator guided design of a new tram line:
Up to six indicators can be configured in the dashboard. For this example four indicators were chosen: 1) Avg. commuting time to jobs 2) Avg. travel time to train station 3) total population in the catchment area of tram stations 4) the relative difference between mean and median of indicator three. Low values of the latter indicate a more evenly distribution of potential passengers between stations. Values can be displayed for the whole city and individually for defined sub-areas.
To compute an overall score, weights and bounds have to be defined for each indicator (this can be done in the set weights tab). For instance, indicator four is only of secondary importance and thus could be weighted with a low or very low coefficient. Bounds define within which range indicator scores high or low. For instance, if planners consider an avg. Commuting time of say 20m as well ideal given the cities structure and an avg. Commuting time of 60 min unacceptable, then the bounds could be set to 20 and 60 min to reflect this viewpoint. All in all, defining a function for the overall score can be a quite complex task that usually prone to be biased in one or other direction.
Costs: costs are a simple linear function of the length of tram tracks and the number of stations multiplied with avg. costs per meter of tracks and the number of stations. When "relative“ costs are chosen the program calculates the costs in euro associated with one changing the indicator by one percentage point (relative to a defined status quo or reference scenario). Screenshot of the tool; Source: Author

In case an average cost was defined (e.g., per meter/square-meter/point), the estimated cost affiliated with each solution is shown and tracked as well. Again, relative values can be shown, more precisely, the monetary costs per increase or decrease of one percentage point of the indicator compared to a reference scenario.
Fig. 32 explains the interface in detail. When fully expanded, the interface might look intimidating at first sight. Therefore, charts can be hidden to simplify the visual experience. Fig. 31 shows the same setup, with most charts hidden.

Mapping

In general, the current state of indicators can be mapped/visualized on the analysis grid, POIs, or street segments (one at a time for each). Furthermore, there is the option to map the difference instead of the absolute value of an indicator (relative to the initial scenario/status-quo).

Automatic optimization

As (a limited set of) indicators are computed on the fly, it would also be inviting to employ genetic-optimization algorithms e.g., for allocating several parks or schools minimizing commuting time to them. Grasshopper comes with two of these algorithms, while plugins such as DecodingSpaces offer additional functionality for the generation and optimization for e.g., street networks. One can also think of combining both approaches: indicator- guided manual design and computational optimization (see Fig. 33).

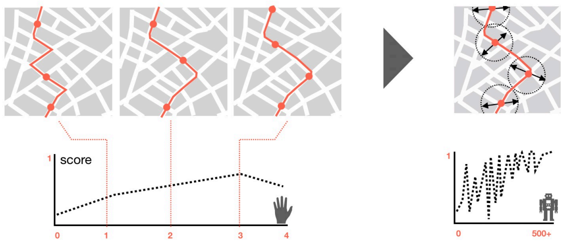


Fig. 33: Indicator guided manual design and additional GA optimization. Concept. Source: Author

Export Scenarios

When all changes have been set, or the user wishes to take the current state of the design process for in-depth investigation, she can click on the export scenario. All indicators will be computed twice, once for the pedestrian/ transit network and once for the car network. The full calculation of betweenness and closeness values is triggered separately as it can take a long time. Calculated values are either stored at the analysis grid, the street segments, or the points of interest. Additionally, the network graphs are exported and can thus be employed in the evaluation-mode for queries or even the creation of additional indicators. Finally, all data is saved in a .ghdata file that can be read by Grasshopper.

The results are saved on the analysis grid, POI locations (points), or the street segments (lines). Therefore a scenario can easily be exported into formats like .shp for GIS software

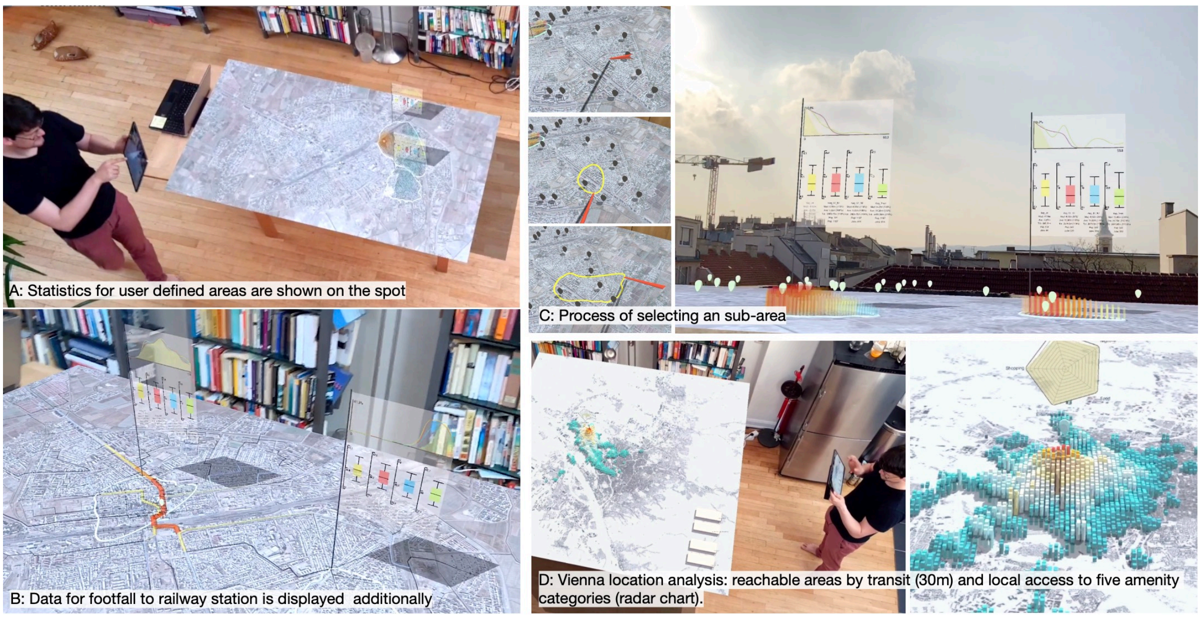


Fig 34: The toolbox in an Augmented Reality environment. Demonstration of an interactive and collaborative workflow within an Augmented Reality environment using Rhino Grasshopper together with the plugin Fologram. Two models are shown: 1) A network-gravity model showcasing accessibility analysis for the city of Kagan(A-C). 2) A network model for location analysis in Vienna. Indicators for reachability and local access to amenities are visualized and can be explored interactively (D). Multiple users can interact with these models by defining areas (draw polylines) clicking buttons or placing points directly via their mobile devices or the HoloLens. Additionally, the toolboxes UI on the computer hosting the models can be used to change parameters and display additional information.
A video where two users interact with the model at the same time along can be watched here: **https://vimeo.com/361153424**

such as QGIS, or as GeoJson, which works great with web mapping services like MapBox. Indicator values need to be saved as attributes on the adjacent geometries.

Moreover, using the plugin Fologram, the data of a scenario can also be explored within an augmented reality environment, using mobile devices or the Microsoft HoloLens. Geometries are directly streamed from rhino to connected devices. Basic interactions are also possible while calculations are done in Grasshopper, and results returned to the devices. See Fig. 34 for two examples of Vienna and Kagan where an early version of the toolbox has been used for.

3.4.3 Evaluation-mode

The evaluation-mode runs in a separate Grasshopper definition that exclusively relies on previously calculated scenarios, saved as .ghdata files as inputs.
Up to four of these scenarios can be loaded at the same time and compared to one another. Next, to the interface, two Rhino layers are (optionally) needed: one to (re-) define sub-areas and another one to define starting points to calculate time radii through the networks as one input option for the DynamicArea option.

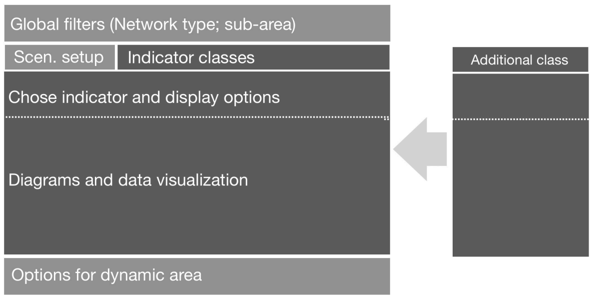


Fig. 35: Interface structure evaluation-mode. The modular structure allows to extend the interface with additional tabs and indicators. Source: Author

The core functionality of this mode is filtering, structuring, and comparing indicators between scenarios. Fig. 35 provides an overview of the user interface structure. On top, always visible, are basic filters for the Network type (pedestrian/ transit; car) and (sub) areas, indicators should be computed for.
Below are several tabs: most importantly for indicators or indicator categories. In case a project requires additional evaluation criteria, a new tab can be attached to the interface with ease.

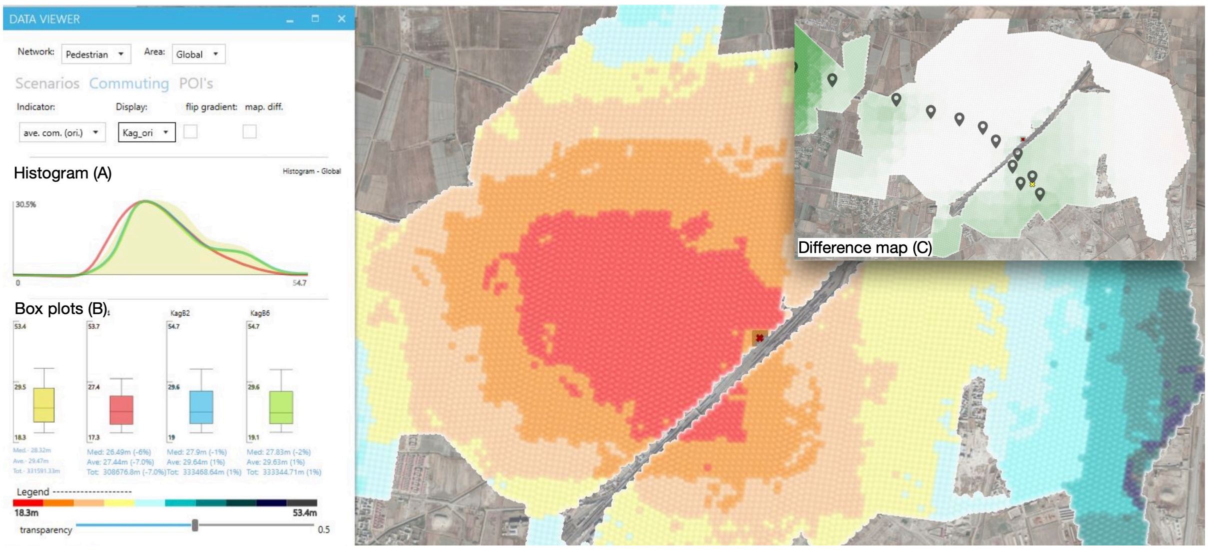


Fig. 36: Dashboard and data visualization for analysis grid. Dashboard for the analysis grid class. Left: Histogram and box plots with indicated relative difference to the reference scenario's values (displayed in yellow). The other scenarios are marked by colors. Right: data visualized by a color gradient (avg. commuting time to work by walking). Corner: difference map between the reference scenario (without a tram) and a scenario with a tram line (avg. commuting time to work). Source: Author

At the top of the dashboard are more options to select the indicator and scenario from. The latter defines which scenario is used to be displayed and mapped to the canvas. Beneath, charts, and statistical indicators can be found. They summarize the data distribution and the difference of each scenario to the reference case. Below the indicator tabs, options to define the dynamic area can be found. Below, some screenshots and examples illustrating the structure and user-interaction with the interface for each of the indicators calculated for, are presented: 1) the analysis grid 2) segments/network and 3) Points of interests.

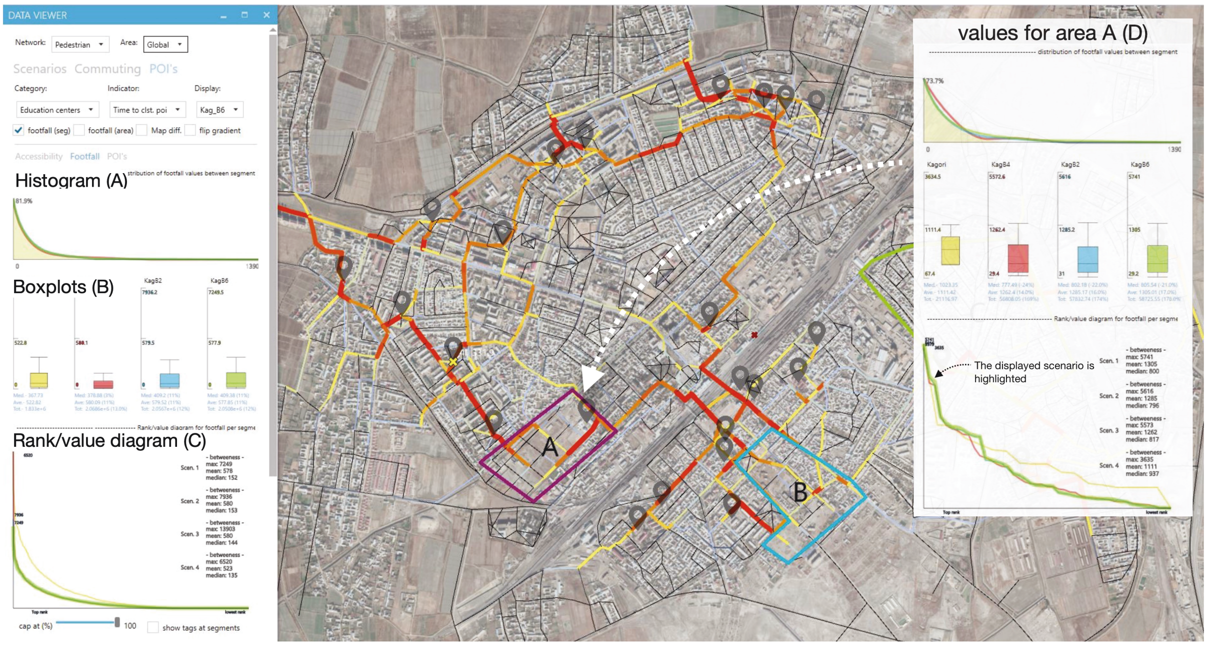


Fig. 37: Dashboard and data visualization for segment (1). Dashboard for the **Segment or network class**. Right: Histogram and box plots work similar as described before. A rank/value diagram is added. Left: Footfall to the POI category education are mapped. Corner: Only values for the sub-area A are used to compute the statistics. This is done for each scenario individually and allows for comparison of filters areas across scenarios. Source: Author

Fig. 36 demonstrates the dashboard for Indicators computed on the analysis grid. Values with spatial dimensions can be mapped to the canvas. Histograms (A) and box plots give an overview of the distribution of a chosen indicator (here avg. commuting time to jobs via pedestrian network). Each box plot and histogram represents one scenario. Optionally a difference map between the reference scenario and one selected scenario can be displayed. The difference map in C shows expected changes in case a tram line is added to the city. The black pins represent the tram's stations.

The dashboard for segments or networks is shown in Fig. 37 and 38. Values for footfall/ betweenness are visualized using three types of diagrams: Histogram (A), box plots (B), rank/value diagram (C). The latter hints on how evenly footfall to e.g., the railway station is distributed over the network: the less the plotted curve is bent toward the bottom-left cor-

Fig. 38: Dashboard and data visualization for segments (2). Map A shows the difference map between the reference scenario and a selected one. Moreover actual segment values can be directly displayed on the canvas. Screenshots of the tool box. Source: Author

ner, the more evenly footfall is distributed. The yellow-colored surface/ curves represent the reference scenario, while the colored lines represent alternative scenarios. Again, user-defined sub-areas can be singled out and looked at in detail (D). Segment values can be mapped by color and line width.

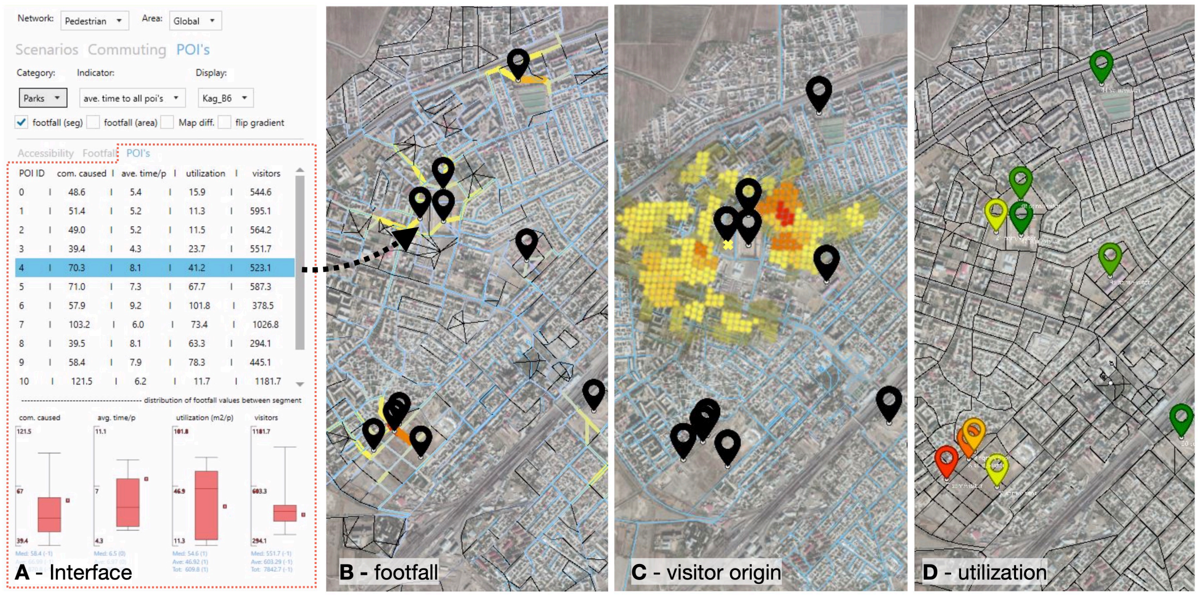


Fig. 39: Dashboard and data visualization for POI class. Dashboard for the POI class (A). The table presents individual POI specific statistics such as visitor numbers, utilization, avg. travel time by visitors and total commuting time caused. Any POI can be selected from the table and is then highlighted in the canvas. The Box plots below show the data distribution for each of the four indicators as well as the median, average and total values - useful metrics for the optimization of both public and commercial entities's location choices. B - D show the mapping options. Source: Author

Fig. 39 A shows the dashboard for POI. For each POI in a category (e.g., shopping centers), four indicators are computed: The number of visitors, the sum of their travel time, the average travel time, and the utilization or use-pressure (available square meter per visitor (FAR/ visitors, if defined)). The box plots provide an overview of the distribution of these values while the table above deals with each POI individually. When clicking on an item in the table, the selected POI will be marked on the map. Optionally the origin of its visitors can be shown as well (C). Additionally, the calculated footfall to the POI can be visualized (B). In image D, the indicator utilization is represented by coloring the marked POI locations with a gradient.

The evaluation-mode is streamlined to compare scenarios to one another, while the functionality of defining sub-areas through filtering provides solid flexibility and rich possibilities to explore pre-calculated data sets.

Next, the functionality is further demonstrated with a short case study of determining the preferable location for a second railway overpass in the city of Kagan, Uzbekistan. The analysis was part of a joint project between the World Bank, Superwien architects, and the Austrian Institute of Technology carried out in 2019.

3.5 Digression: A second pedestrian railway crossing in Kagan

This section evaluates six options for a second pedestrian railway overpass (Fig. 40). The results are summarized in Table 4. In terms of indicators, a focus on overall connectivity and accessibility was given. After a first round, three candidates could be eliminated. For the remaining ones (A, C, and D), further analysis of expected impacts on pedestrian flows among others was conducted.

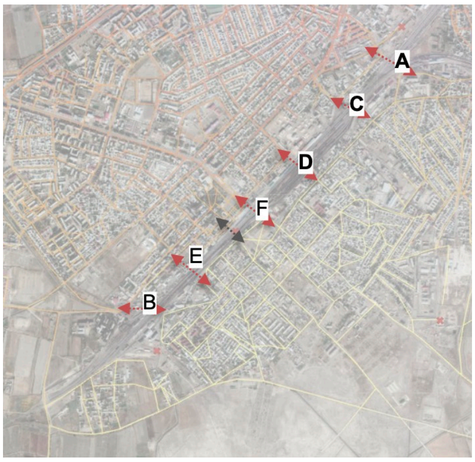


Fig. 40: Possible locations for a second bridge. Source: Author

The calculations are based on the same graph-gravity model that is part of the toolbox. Its parameters were estimated based on feedback from the workshops with experts conducted in Kagan and literature.

Most Indicators make a statement on the overall change compared to the status quo, taking the whole city into account. Local impacts can be much stronger (e.g., a new bridge can cause improved access from one area to all others by 23% and a particular area, directly on the other side of the bridge, by hundreds of percentages while the overall improvement can be as little as 3%).

If the amplitude of change of the median is higher than in the mean, it indicates that the gained access/ benefit is spatially more evenly distributed.

Indicator/ Location	**Location A**	Location B	**Location C**	**Location D**	Location E	Location F	Indicator description
Population within 10m walk	3300	4500	3500	4000	3700	4800	Number of people living in direct proximity to the bridge (Assuming a total population of 58 000)
Jobs within 10m walk	980	880	1200	1400	1400	1800	Number of jobs in direct proximity to the bridge (assuming 12000 total jobs in Kagan)
Access to people (improvement in %)	**3.1%** (mean) **22%** (max)	**0.62%** (mean) **11%** (max)	**3.8%** (mean) **20%** (max)	**3.2%** (mean) **15%** (max)	**0.23%** (mean) **4%** (max)	**0.9%** (mean) **3%** (max)	Sum of travel time to all other areas weighted by population. A measure for connectivity and interaction potential.
Access to jobs (improvement in %)	**3.2%** (mean) **22%** (max)	**0.5%** (mean) **9%** (max)	**4.5%** mean) **22%** (max)	**3.5%** (mean) **16%** (max)	**0.25%** (mean) **3%** (max)	**0.91%** (mean) **3%** (max)	Sum of travel time to all other areas weighted by jobs. A measure for connectivity and interaction potential.
Avg. commuting time to jobs (walking) (improvement in %)	**2%** (mean)	**1%** (mean)	**2%** (mean)	**2%** (mean)	**0%** (mean)	**0%** (mean)	Based on previously estimated origin/destination matrix from population to jobs.
Avg. job opportunity index (walking) (improvement in %)	**2%** (mean) **3%** (median)	**1 %** (mean) **0%** (median)	**3 %** (mean) **4%** (median)	**2 %** (mean) **3%** (median)	**0 %** (mean) **0%** (median)	**1 %** (mean) **1%** (median)	The indicator is calculated for each spatial unit: Jobs add to the score with a inversely proportional function to their distance.
Walking time to all shopping centers (improvement in %)	**3%** (mean) **5%** (median)	**1%** (mean) **1%** (median)	**3%** (mean) **5%** (median)	**2%** (mean) **4%** (median)	**0%** (mean) **0%** (median)	**2%** (mean) **2%** (median)	Average walking time to all shopping centers (POI)
Walking time to all parks (improvement in %)	**2%** (mean) **3%** (median)	**1%** (mean) **3%** (median)	**2%** (mean) **2%** (median)	**2%** (mean) **2%** (median)	**0%** (mean) **1%** (median)	**1%** (mean) **1%** (median)	Average walking time to all parks (POI)
Walking time to Railway-station (improvement in %)	**0%** (mean) **0%** (median)	**0%** (mean) **0%** (median)	**0%** (mean) **0%** (median)	**0%** (mean) **1%** (median)	**0%** (mean) **0%** (median)	**1%** (mean) **1%** (median)	Average walking time to railway station (POI)
Local integration, connectivity r1600 meter (description of impact)	High	Low	High	Medium	Low	Medium	Closeness centrality for radius 1600m. Gives a sense and how the connectivity of neighborhoods increase locally
Local integration, pedestrian flows r1600 meter			See images				Betweenness calculation for radius 1600m. Gives a sense on which roads gain and loose in footfall potential and importance.

Table 4: Overview of indicators and results for Kagan. In green: short-list of locations for further analysis. Option C ranks first in six out of ten indicators and therefore outperforms both A and D in the quantitative analysis. Source: Author.

Rank　Rank 2 or 3

The most relevant indicators are access to people and access to jobs (giving a good proxy for the overall accessibility to places of potential interest) followed by local integration, as it can be interpreted as a proxy for the reach or size of a neighborhood or as an indicator of the spatial integration of areas. Additionally, betweenness centrality values were calculated and reviewed (see Fig. 45) to investigate which roads gain and lose in significance. Below some details on the three best performing options are given:

Impact on access to people and jobs (Fig. 41): Option A delivers a solid improvement of the current situation. However, looking at the difference map, it gets apparent that the proximity to the town's border causes lost potential. Compared to A and D, location C performs between 20% and 45% better. Option D comes with a significant improvement, as well. The low maxima values indicate the connectedness of already well-connected areas is further improved, which tends to strengthen the existent center.

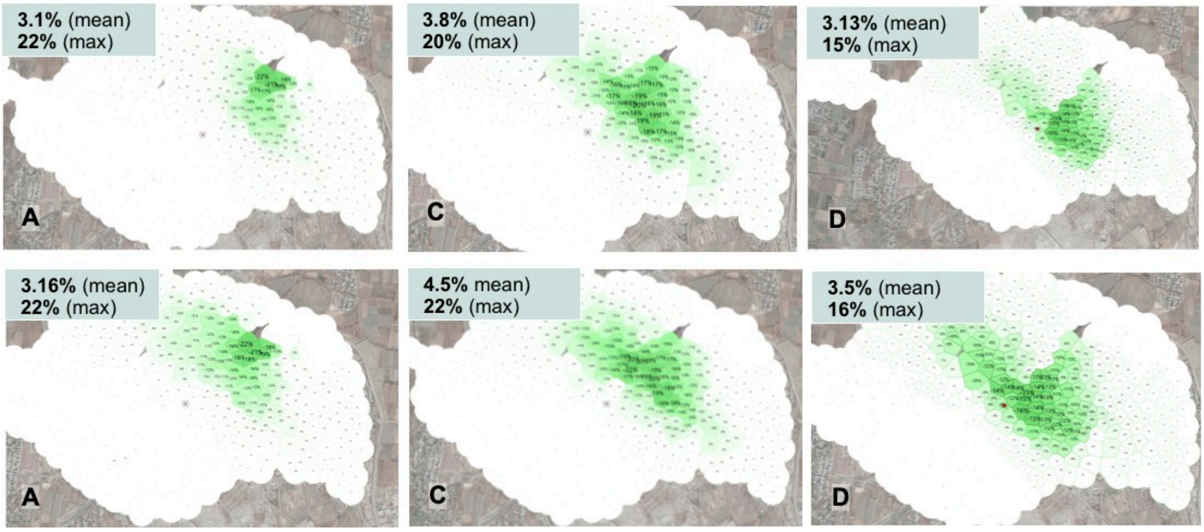


Fig. 41: Impact on connectivity. Top: access to people; Bottom: access to jobs. Source: Author

Impact on footfall to major shopping destinations: A new bridge at location A has almost no impact on the footfall potentials to shopping centers. Location C would have a significant effect on the expected footfall to shopping destinations. Additionally, it serves as a direct link to the only larger commercial entity in the south-east of Kagan. Thus, potentially streng-then it. Option D has a similarly significant effect on the estimated pedestrian flows. The second bridge improves access to the south-eastern commercial unit slightly and by-passes some pedestrian traffic away from the town's center (Fig. 42).

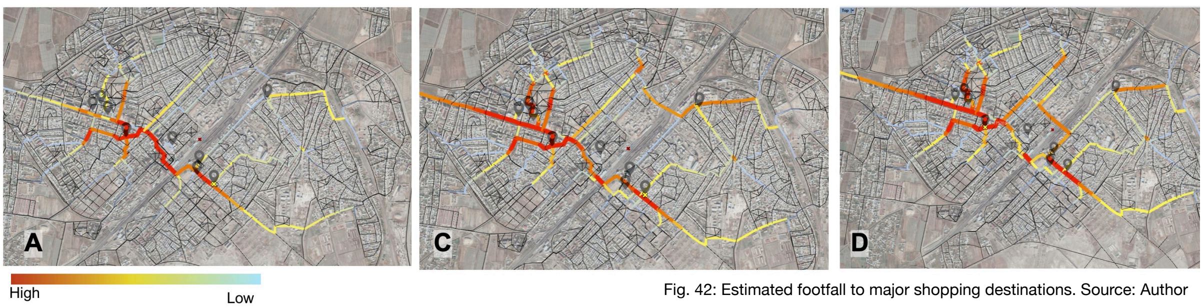


Fig. 42: Estimated footfall to major shopping destinations. Source: Author

Impact on the reachability of shopping centers: A, C and D have the most substantial impact (mean of -3%, median of -5%, compared to the status-quo). Moreover, the histograms visualize that most areas with a previous travel time between 32 min and 44 min to all centers benefit from a second bridge (Fig. 43).

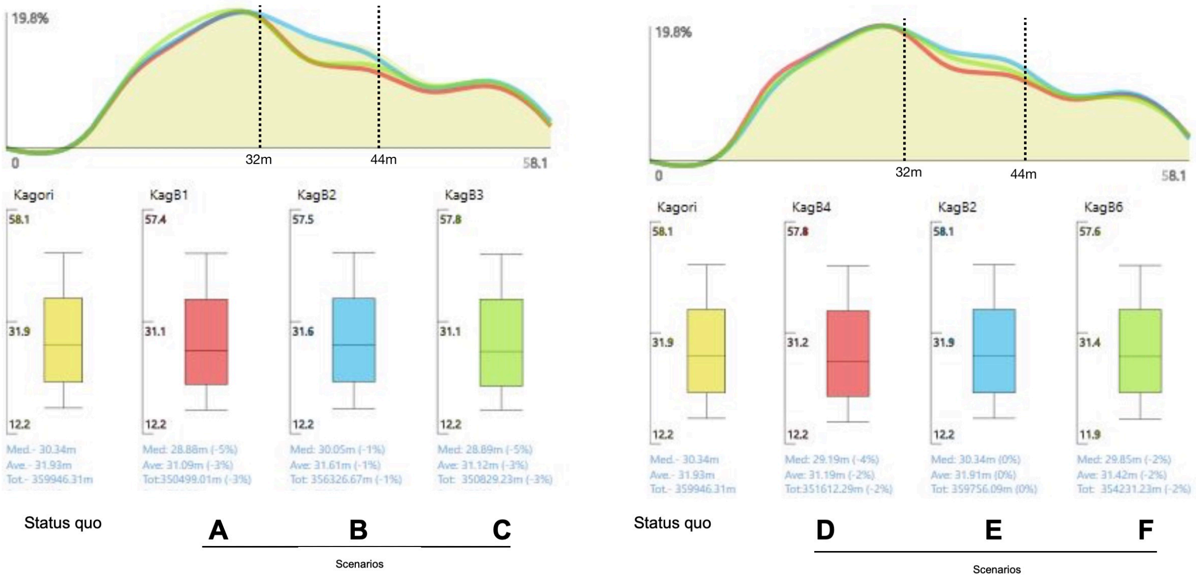


Fig. 43: Total travel time to all shopping centers. Source: Author

Impact on local connectivity and neighborhood cohesion: Large parts of the areas benefiting from location (A) in terms of increased local connectivity are low-density or non-residential. Moreover, as the bridge is closer to the north-eastern border, some potential gets lost, resulting in relatively few people benefiting from this option (Fig. 44).
Option C has the largest overall impact of increasing local connectivity. Moreover, it reaches out to all sides strengthening the bridge's potential as a local hub. A bridge at location D would have a significant positive effect on a moderately large area. However, mostly already well-connected neighborhoods will benefit from this constellation.

Fig. 44: Areas experiencing major improvements in local connectivity. Computed via closeness centrality. Source: Author

Impact on pedestrian frequency potentials: The proposal A affects the pedestrian traffic on the original bridge least (-19%), while the northern railway crossing is estimated to be frequented by -50%, which hints that movement currents between both sides of the city are evenly served by this constellation (see Fig. 45).
Compared to option A, C is better integrated into residential quarters on both the north and the south side. Moreover, the fact that the frequency on the original bridge is decreased by 33% hints at an even greater improvement of origin and destinations relationships across the railway in terms of connectivity.

Here, the second bridge is in stronger competition with the original one. According to the estimation, the first bridge would loose up to 60% of visitors to the new one. In the likely case, that capacity is not a significant issue. his fact would render option D rather ineffective.

Fig. 45: Estimated changes in pedestrian frequency potential for the options A,C and D. Source: Author

Conclusion: All in all, option C clearly outperforms option A, while in terms of numbers also dominating D. The final decision between C in D is a matter of normative standpoints: If spatial equality and a more evenly distributed access to all citizens is prioritized then option C appears optimal. In case strengthening the center and perhaps fully integrate parts of it into old Kagan is aimed at, location D might be the better option, despite its lower impact on the overall connectivity.

Discussion: Overall, the toolbox proved useful in facilitating an evidence-based discussion and decision process on the ideal location for the second railway overpass and was received positively from the project partners. The combination of comparing multiple scenarios against one reference scenario at the same time with spatial filtering options turned-out to be a robust framework to compare alternatives for one project but also to test for a ideal combination of multiple projects: As part of the same project, the urban design studio was asked to develop around six interventions for the town, which were then split into eight components (e.g.adding new parks, renovating the bazaar, adding new residential areas, adding a tram line, etc.). These components had to be ranked in a cost-benefit (or cost-impact) analysis in order to bundle 3 - 6 projects together that are to be financed. Here, the tool was successfully used to find an effective combination of components that maximized their overall impact.

3.6 Reflection

IIt is important to emphasize that the toolbox at no point claims to cover all issues related to an essential for planning and monitoring urban systems. Thus, it can only contribute to the fields it makes dedicated statements on. Moreover, its accuracy crucially depends on the data quality used as input and the calibration of its parameters. As both can be challenging, the toolbox can be seen as a planning and design decision support system - as one additional factor to base choices on.
However, even in cases where the data quality is questionable, the toolbox could still be used to compare relative differences between scenarios or in order to better understand and discuss the underlying dynamics of urban systems.

Looking a little forward, to the near future, one can expect that the application of Deep-learning and recent advancement in remote sensing and object detection from satellite imagery allows for semi-automatic generation of most of the input data the toolbox requires. This would expand the tool's capability of monitoring by a significant degree. Features extractable of aerial images include street-graph, building footprints, and (a set of) classes, as well as land uses such as residential, industrial, and so on. Fig. 45 provides some examples of recent achievements in this field: A) shows the extracted building footprints, B) waterbodies, C) a street-graph, and D) an example of a model that recognizes multiple classes in one image (houses, highways, sidewalks, etc.). Besides, also smaller objects such as cars can be recognized. Fig. 46 describes the process in more detail: For the LNA Case study, it was planned to automatize the counting of cars (as a proxy for activity levels) using the YOLOv3 neural network that is able to spot multiple objects with bounding boxes in images (Redmon 2018). A Grasshopper definition was set-up to quickly label additional data for the training of the model within the Rhino viewport. Due to time constraints and a lack of compatible hardware, the development, unfortunately, had to be stopped after the stage of training data preparation.

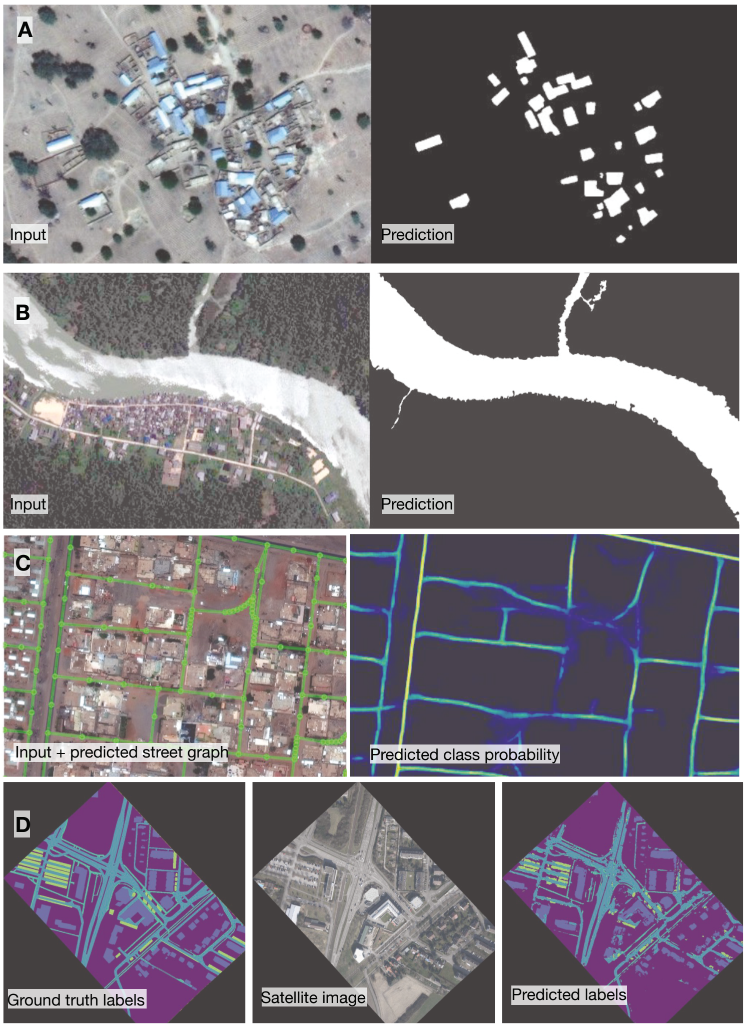


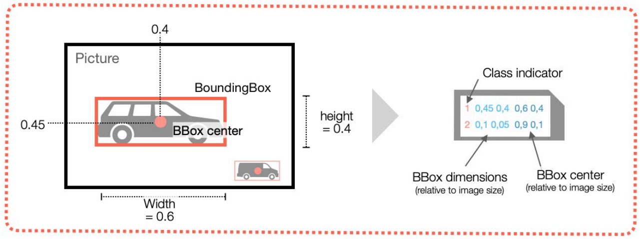
Fig. 45: Deep-learning based feature extraction from satellite imagery. Source: A and B: Nowaczynski 2017; C: Van Etten 2018; D: Sprengel 2017.

For an application in existing towns and settlements, the integration of additional data should be considered as it has the potential to extend the models' accuracy and scope significantly. The ongoing research project "An indicator system for monitoring human settlements in the environment of new data "(新数据环境下的人居环境监测指标体系) led by Long Ying (see Long et al. 2017) gives a profound overview of (open) data sources and their application that can be tapped to support urban planning and simulation efforts. Although focusing on China, they can be transferred to other contexts. Another collection on the usage of emerging big data can be found in the book Decoding the City: Urbanism in the Age of Big Data (Offenhuber 2015). Proposed sources include the crawling of websites to gather data on property prices and rents, using web-map API's to get data on traffic, real-time activity levels (via active Baidu maps users in an area), or anonymized cellphone location data among many more.ny more.

Applying neural networks to localization and count cars on satellite images using the yolo v3 architecture and grasshopper for data preparation

Fig. 46: Data preparation of the yoloV3 model for the detection of cars

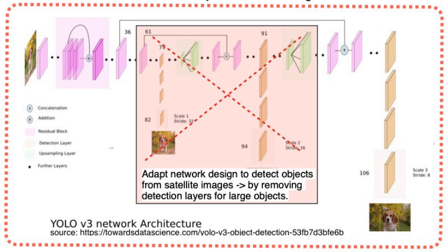
Data preparation (finished)



BBox can either be drawn with a rectangle or, in case only a point is provided, set as a default sized one.

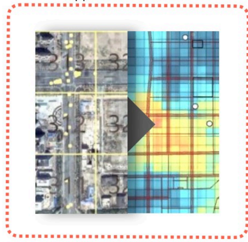
For each image tile a .txt file is created that carries information for each object that was marked within.

Model adjustment & training



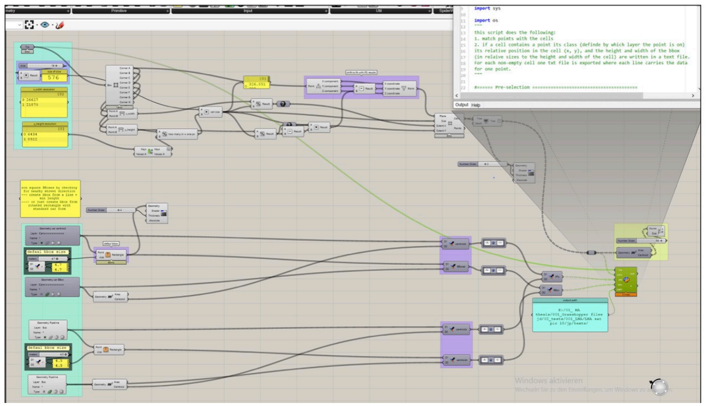
Feed images and .txt files to YOLO v3 for training.

Application/ detection



Feed new satellite images, detect, mark and count cars on them

The Grasshopper definition for marking cars and writing the txt file:



Data Sources

1. New satellite imagery from the case-study area, manually annotated in Rhino/Grasshopper (the background sat-image was splitter into cells, cars were marked with points, a default sized BBox was casted around each point, its relative coordinates and position were saved in a text file, one for each non-empty cell (using a python script)
2. A large existing dataset COWC („Cars Overhead With Context") was used and transformed with a definition in grasshopper: For each satellite image an additional .png with colored pixels only at the position of a car is supplied with the dataset. The definition detects the pixels, its color (defines the class/type of the car) and then writes the .txt file in the needed format (see above).

Overall more than 5000 marked cars were obtained. Before „training" the images should all be re-scaled to the same resolution the test images are expected to be. Adding training samples from the case study area into the pre-trained model expectantly increases the models accuracy.

4. Application: The Lanzhou New Area

4.1 Introduction

The Lanzhou National Level New Area (LNA) is a new town project in Gansu province, located roughly 60 kilometers north of the provincial capital Lanzhou (1.8 million inhabitants) (Fig. 47). It aims to inhabit a population of one million by 2030 and serve as a major industry and logistic hub in western China. It is part of National Level New Area program (NLNA)[10], the Go-west strategy[11], and also hosts a model project for the National Plan for New-Type-Urbanization (see CAUPD 2015, LNAAC 2015). Finally, being located at the hot-spot of the ancient silk-road bridging east and west and, after Urumqi, being the most western city with solid infrastructure and logistics capacities, the Lanzhou and the LNA are also prominently positioned in the *One Belt, one Road* meta strategy launched by the central government in 2013.[12]

Fig. 47: The LNA within chinas urban system. The red star makes the position of the LNA. Red circles represent other National Level New Areas. Designated city clusters are encircled in yellow. Source: Author with data from State Council 2010.

Despite the many programs, attached subsidies and preferential policies, the Lanzhou New Area poses a potentially interesting example of a nationally prominent, large- scale urban development project that faces a great risk of not reaching its original development goals and target population in the foreseeable future (NDRC 2017, Ding 2019).

The LNA became the fifth member of the National Level New Area program in 2012. The first NLNA was launched in 1992 (Shanghai Pudong) and 2006 (Tianjin Binhai). They were extraordinary successful with annual growth rates (GDP) averaging in over 30% for the first 15 years of development (NBS 2018). Both are in direct proximity to global trading hubs and located in the highly developed coastal areas of China, well integrated to major urban agglomerations. Despite striking differences (weak economic performance of Gansu, no access to ports, high distance to a relatively small mother city), authorities anticipated experiencing similar growth rates for the Lanzhou New Area as have been seen for the for its early counterparts (Xu 2017).

---

[10] National Level New Areas (NLNA, 国家级新区) are large-scale, mixed-used urban development projects that can bee regarded as an increasingly important strategic tool of the Chinese state to induce new impulses on economic growth, development, policy innovation and reform implementation and guide development directions on the regional level. NLNA are entities with a high administrative status that usually enjoy preferential policies granted by the central government. In addition they are under direct supervision by the latter. The first NLNA was Shanghai Pudong (1992), followed by Tianjin Binhai (2006), they can be interpreted as large and more mixed use special economic zones. After 2010 (4) and especially 2014 (12), sixteen more National Level New Areas were launched, thereby indicating an increased importance attributed to this particular tool of development and spatial policy on the national scale (NDRC 2015; Wu et al. 2015).

[11] The go-west strategy（西部大开发) was originally launched in 1999. The term subsumes various nationwide efforts to support western provinces in catching-up with the economically over-performing costal areas and facilitate a spatially more even distribution of economic activity (see Roland Berger 2009).

[12] See Shih et al. 2014 for a brief overview on the *One Belt, One Road* program.

Fig. 48: Impressions of the Lanzhou New Area.
The map on the righthand sight shows the black plan of both the LNA and Lanzhou. Source: Author with image data from google. Photos: http://gansudaily.com.cn.

Therefore, constructions were executed rapidly, work on the infrastructures (such as roads, sewage system etc.) is already completed for large parts of the 1200 km2 large area (see Fig. 48 - 50). Moreover, construction on residential and industrial areas has begun in a top-down approach which led to dispersed patterns of land uses: projects are scattered within a radius of 15 km. This clearly makes it harder for the new town to develop a working spatial structure and gain the critical mass to enter a self-propelling growth process of attracting people, commerce and jobs. It is, therefore, facing an increased risk to become one of the infamous "ghost towns“ of China (Shepard 2015).
The short time between creating the plan and the start of large-scale infrastructure construction led to problems in adapting the city's spatial-structure to later insights (such as an initially wrongly calculated dominant wind direction, see CAUPD 2015), policy requirements such as the *National Plan of New-type Urbanization*[13] or a much slower than anticipated capability of attracting both population and businesses.

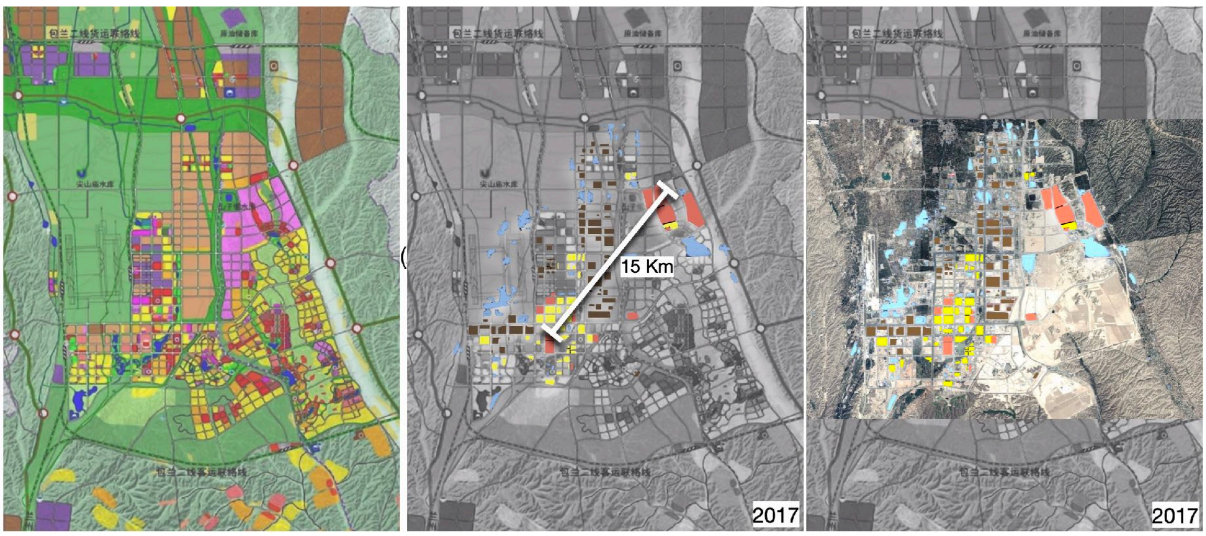


Fig. 49 : Masterplan and current state of development.
Left: Masterplan; middle and right: Developed plots highlighted. Colors indicate land use -yellow: residential; red: commercial and education; brown: industrial; blue: villages. For an overview on the Chinese planning system and the role of masterplans see Wu 2015. Source: Author

[13] The National Plan of New Type Urbanization, (国家新型城镇化规划 (2014-2020),formulates new requirements toward urban development in terms of social issues (e.g. integration of migrant workers), macro-level urban systems, function mix in cities, urban design and more. Overall, these policies try to facilitate a departure from a *growth first, at any price* approach toward a more socially cohesive, sustainable and land efficient urbanization *for* the people (人的城镇化) (State Council 2014).

After more than seven years in development, the discrepancy between the ambitious goals and reality are steadily growing. Its population reached around 50% of the target for 2015 (mostly relocated farmers); the economic output 25% (Xu 2017). Despite the fact that seven years is a concise time horizon to look at urban growth, the comparison with the trajectories of other NLNA, emphasizes the development problems of the LNA: for instance, the investment density per hectare is at 46 Mio. RMB compared to 72 Mio. in other NLNA located in Western China, indicating a low land use efficiency (Xu 2017). The Real estate market, as of now, primary depends on relocated farmers and workers. Despite low prices (avg. of 4000 RMB per square meter as compared to 10000 in the mother city Lanzhou), the average monthly transaction rate of properties square meters is 15 times lower than in Lanzhou (48 thousand vs. 700 thousand) (Song 2018). Additionally, the vast development led to a significant mismatch between infrastructure constructed and actual people using it: The estimated road-surface for 2017 numbers to 165 square meters per potential habitant (if all apartment units would be fully inhabited). This number would even be much higher when calculating with the actual population. In comparison, 65 square meter road surface per capita are calculated in case the masterplan is fully implemented, and all buildings inhabited (numbers were estimated using the toolbox).

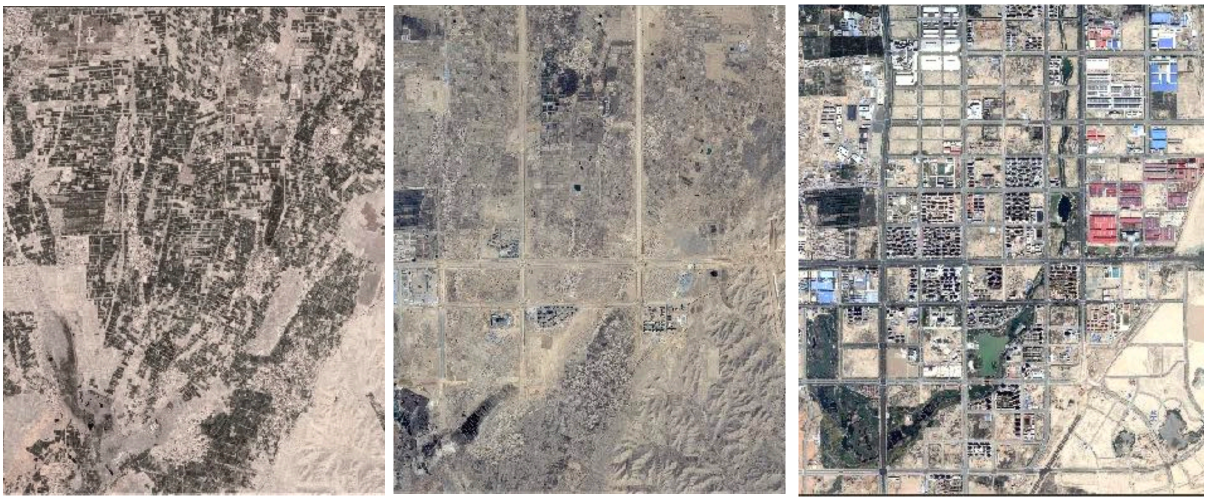


Fig. 50: Satellite images showing the LNA's development. Left: 2011; middle: 2012; right: 2017. Source: google

For these reasons, the LNA offers an interesting case to conduct the outline analysis and test the framework: As a completely new development, almost in the nowhere, it provides a rather pure example of new town development, relatively remote from other forces that potentially drive urban development. Thus, cause-effect relationships and changed dynamics induced through the masterplan are easier to isolate and interpret. Moreover, toolboxes like this might also be applied for the analysis of planning proposals (and evaluation of their implementation) toward their coherence to policy targets formulated in the Plan of New-type Urbanization. For instance, the avoidance of developing mono-functional areas or less than 100 square meters of built-up area per habitant in urban areas (State Council 2014).

4.2 Data preparation and setup

Overall, there is not much data readily available for the Lanzhou new area. For the following analysis, mostly high-resolution satellite imagery (obtained from google) and published documents on masterplans were used.

Satellite images were used to count driving, and parking cars, for constructing the street networks and to derive block typologies (see below). From the masterplan information on land uses (block level), amenity locations and the street network were obtained. The figures 51 - 53 illustrate the process of how socio-demographic attributes for land uses categories were derived: First, empirically data on population and job densities for 32 land use categories was obtained from a traffic impact estimation manual from the city of Suzhou (PBS, 2015). Afterward, the built-up volumes of several already developed blocks of the Lanzhou new area were roughly modeled (see Fig. 53) to adopt the numbers to the local context. Finally, for the main land use categories (used in the LNA masterplan), average density values for residents and jobs per square Kilometers were calculated, which later serve as weights for the calculation of indicators. The results are presented in table 5.

For the setup of the street network graph, five road categories were distinguished (see table 6). Primary roads are modeled as two parallel running one-way streets to account for their

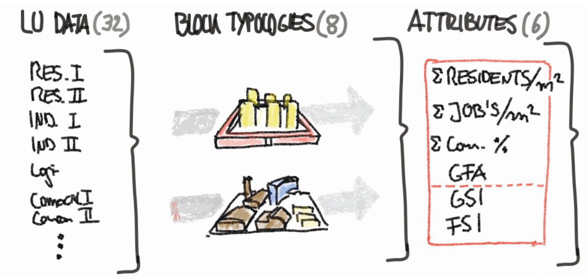


Fig. 51: Procedure of deriving land use typologies. Source: Author

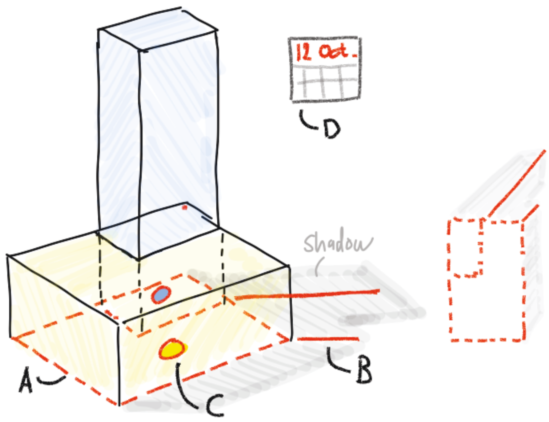


Fig. 52: Estimating GFA and land use shares on the block level.
The first step is to obtain satellite imagery of the site with a know date of recording (D). Next, the building footprints need to be redrawn by polylines (A). Then, a single line has to be drawn following the shadow casted by the building (B), it is important that the correct angle is used. With these two inputs the building volumes are calculated. If volumes are intersecting, the building with the larger footprint culls the other. Finally, land uses can be assign by placing points in the drawn shapes (c). Implemented in *Rhino/Grasshopper.* Source: Author

	pop. / km2	Jobs / km2	FSI	GSI
Village	2000	250	0,61	0,6
Residential block	30000	3000	1,5	0,22
Mixed use block	15000	8000	1,7	0,25
Commercial block	6000	15000	1,7	0,34
Industrial block	0	8000	0,45	0,34
Light industry/ logistics	0	5000	0,45	0,43
Education and research	5000	10000	0,7	0,15
Green spaces	0	80	0,05	0,05

Table 5: Socio-economic statistics derived for land use categories. Source: Author

Parameter	Value
Speed highway (m/s)	32
Speed primary (m/s)	17
Speed secondary (m/s)	13
Speed tertiary (m/s)	8,2
Speed pedestrian (m/s)	1,4
Speed tram line (m/s)	10
Avg. waiting time tram (s)	300
Park (ß1; ß2;a)	0,003/0,002/0
Schools (ß1; ß2; a)	0,003/0/0
Jobs (ß1; ß2; a)	0,0004/0/0

Table 6: Parameter values used for the LNA case-study. Beta values are scaled to time units (s). Source: Author

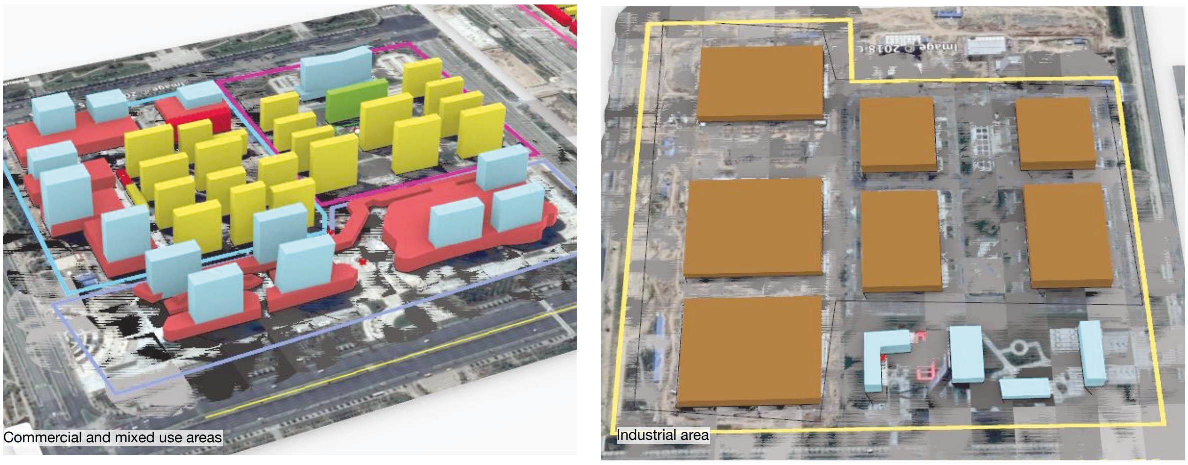


Fig. 53: Examples for existing blocks in the Lanzhou new area.
Colors represent assigned land uses. Yellow: residential; red: commercial; blue: office. The blocks were assigned to categories (e.g. Residential, mixed use, industrial). The categories average values on residents and jobs per square meter among others were then employed as weights in the model. Source: Author

size (between 40 and 80 meters in width) and poor crossing possibilities.
Unfortunately, it was not possible to obtain detailed plans of the transit system. Instead, only one light rail line depicted in the masterplan was included in the transit graph. The value of the road network related parameters, along with others is depicted in table 6.

4.3 Additional Indicators

Arguably each planning or analysis undertaking is unique and thus demands an adopted quantification and evaluation system. Below, an indicator system is presented that aims to capture the development dynamics of the Lanzhou New Area on the one hand, and on the other allows to quantify, compare and evaluate different spatial configurations. All analyses operate, relatively aggregated, on the city level, not allowing for e.g. detailed statements on issues regarding neighborhood levels.
Besides, the reviewed call toward a more evidence-based planning approach, along with the constant monitoring of a city's development, further guided the formation of the indicators. Therefore, special attention was given to the attempt of capturing the frictional relationship between bottom-up and top-down driven developments as well as the discussed degree of coherence with the social logic of space. The indicators are subdivided into four meta-categories (infrastructure efficiency, economy, social (urbanity), and coherence of plan and market (the latter is understood as self-organizational processes inherent to cities). It should be noted that at no point it is claimed to cover all relevant aspects. The presented indicators are instead to be seen as an exploratory advance of examining in which direction an evaluation system could be developed, in order to serve the analysis. Moreover, results should be interpreted carefully. Nevertheless, especially different scenarios and time series can yet be analyzed in relative terms within the naturally limited scope of the suggested toolbox and indicator system.

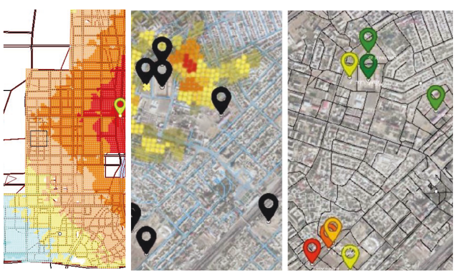


Fig. 54: Indicators of the section *infrastructure efficiency.* From left to right: time to closest facility, visitor origin, utilization/use-pressure. Source: Author

Infrastructure efficiency (Tab. 7): This section covers a few indicators to quantify the costs and benefits (e.g., in monetary terms or land consumption) of infrastructure related aspects. This covers issues such as road surface and cost per capita as well as accessibility to

ID	Spatial resolution	Indicator	Unit	Notes on calculation	Rationale/Interpretation	Reference
1	-	Road surface per capita	Sqm	Road surface / population	Measures the road surface per capita. When tracked over time, the indicator allows to make statements on how effective the cities physical growth is synced with its demographic development.	Rode et al, 2015
2	-	Built-up area per capita	Sqm	GFA / population	A measure of population density. When looked at in time serious the indicator can additionally be interpreted as a proxy of the balance between demand and supply.	NDRC 2015
3	Nodes	Access to parks	Score	See basic indicators	The location of parks and schools are evaluated in terms of visitor numbers, utilization, accessibility and travel times (see table **1** for a detailed description).	
4	Nodes	Access to schools				
5	Nodes	Time to old city (Lanzhou)	Min	Travel time to Lanzhou from area i	Avg, travel time to the next larger city: Lanzhou	Piovani 2018
6	Nodes	Time to Railwaystation stations	Min	Travel time to railway station from area i	Avg, travel time to the railway station	

Table 7: Indicators of the section *infrastructure efficiency.* Source: Author

public amenities (see Fig. 54). For instance, this set of indicators can be used to evaluate whether similar service levels and functionalities are achievable with fewer inputs. Moreover, when looking at time series, it might help to identify inefficiencies in growth trajectories and over- or undersupply of infrastructures.

Economy (Tab. 8): Economic factors considered here are mostly rooted in the literature that looks at the economic potential of cities from a macro perspective. These stem from economies of agglomeration and depend on the effective size of the labor market, particularly. For instance, Bertaud (2018) shows that doubling the size of a talent pool increases the

ID	Spatial resolution	Indicator	Unit	Notes on calculation	Rationale/Interpretation	Reference
7	Distance matrix	Avg, commuting time to jobs	Min	Avg, travel time to jobs for area i , Based on trip distribution from residential areas to jobs via shortest-path,	The average commuting time is an important indicator especially when monitored over time. Where an increase might hint to emerging bottlenecks in the transportation network or sprawling developments. Especially in smaller regions (where travel times are mostly below the typical threshold of one hour) avg, commuting time should be looked at together with the indicator on job opportunity,	Bertaud 2018
8	Distance matrix	Labour-market integration	%	Jobs within one hour (car/ Transit) from area i / all jobs	The effective size of the labor in a city has a significant impact on its overall economic productivity, Bertaud 2018, argues that one hour of commuting is a typical threshold where labor markets start to separate. The measure is a function of density, land use distribution and the transportation network.	Bertaud 2018
9	Distance matrix	Job opportunity	Score	Accessibility to jobs for area i (see table basic indicators)	The sum of this value gives a sense of how effective Residential and work locations are located in space. Regionalized values can be used as a rough proxy for housing prices and attractiveness for residential land uses.	Adoption of Bertaud 2018
10	Distance matrix	Utilized job opportunity potential	Score	Avg, Job opportunity in mono centric city / avg, job opportunity in current spatial configuration	Compares the actual access to jobs with the case of a hypothetical land use distribution in a mono-centric structure: All jobs are in the center and residential uses in outer rings sorted by density (high: center, low: outskirts), The latter represents the theoretically nearly ideal access to jobs (Long et al. 2013**)**. The indicator improves comparability with other cities by taking spatial properties (or limitations) into account.	-
11	Distance matrix	Utilized agglomeration potential	Score	Total of agglomeration index (A): A for location i equals the sum of jobs in location j discounted with a simple gravity function (beta parameter was empirically estimated by Graham et al. 2010).	High levels agglomeration is associated with higher total factor productivity (Graham et al. 2010) and economic potential. This indicator makes a statement on how effective the spatial structure enables agglomeration effects, Similar to indicator 10 the actual situation is compared to a hypothetical, mono-centric scenario with land uses of high job and population densities are allocated most central.	Graham et al, 2010

Table 8: Indicators of the section *Economy.* Source: Author

productivity of a city by 15%. Thus, the performance of large cities largely depends on an efficient transportation system that interconnects residential areas to job centers in a reasonable time. Another dimension looks at the overall size of agglomeration or gravity potential a cities spatial structure realizes. This could be regarded as a relevant measure when looking at regional economics and inter-city competition (see Fig. 55). Looking at time series of e.g. the regional interaction potential, might help to quantify the disadvantages of a scattered development as can be observed for the Lanzhou new area.

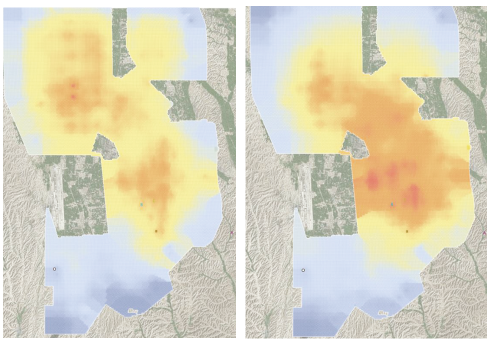


Fig. 55: Indicators of the section *Economy.* Left: Actual job opportunity; right: job opportunity with fictional distribution of jobs and population. The indicator utilized job opportunity describes how close the actual value of *job opportunity* comes to the hypothetical one. See table for more details. Source: Author

Social (urbanity) (Tab. 9): This section picks up the often-referred-to triad of density, diversity and accessibility and their determining effect on urbanity. A term that tries to capture factors that tend to make cities livable, vibrant places of human exchange and creativity as

ID	Spatial resolution	Indicator	Unit	Notes on calculation	Rationale/Interpretation	Reference
12	Distance matrix	Access to global centrality (r = 5000m)	Score	Access to globally highly integrated streets (discounted with gravity function)	A proxy access the spatial distribution of potential (local) centers and access to them, High values indicate that a larger portion of residents has access to centers in walkable distance.	Bielik 2018; Karimi 2014
13	Distance matrix	Access to local centrality (r = 800m)	Score	Access to local highly integrated streets (discounted with gravity function)		
14	Distance matrix	Diversity: Land-use mix	Score	Shannon index of diversity for all LU within 10m walking time, Shannon index for area i: sum (m2 land use x / total m2) * ln(m2 land use x / total m2) * -1	Higher diversity is associated with higher urbanity. The Shannon index is often used to describe species diversity in biology but has also been applied in other fields including urban studies (see Zhang 2017). Diversity peaks when all categories (eg, of land uses) are represented in equal quantity. Here, a diversity value of three can be interpreted that three different land use of equal total GFA can be reached within a ten min. walk.	Zhang 2017; Ye et al. 2013
15	Nodes	Access to population	pop,/km2	Sum of population within 10min travel time	Higher density levels are associated with higher levels of urbanity or spatial capital (Marcus 2007). A higher road network density is often associated with better local accessibility and more active facades etc.	Marcus 2007
16	Nodes	Road network density	Len/km2	Sum of road segment length within 10min travel time		

Table 9: Indicators of the section *Social/urbanity.* Source: Author

well as an attractive destination to move to. It draws on a research branch famously started by Jane Jacobs. More recent scholars emphasize the positive effects on the economic sphere as well (see Hospers 2003; Florida 2004). Diverse urban areas draw more various talents to towns and facilitate the exchange of ideas and knowledge, which can be regarded as the fuel for the more and more knowledge and service-based economy. Fig. 56 shows the indicator *land use diversity.*

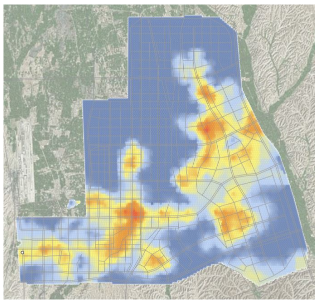


Fig. 56: Indicator for land use diversity. Source: Author

The coherence of plan and market: This section aims to include the friction between "top-down“ planning (master planning) and the social-logic inherent in spatial structures by looking at the matching rate (See Fig. 57). A temporal dimension is added by considering the relationship and time lags between structures being built and activity levels picking up.

ID	Spatial resolution	Indicator	Unit	Notes on calculation	Rationale/Interpretation	Reference
17	Distance matrix	Match rate	Score	Correlation between values for urbanity (indicators of section social) in area i and values for center potential in area i for	In close relation with the match rate / urban maturity indicator from Ye et al, 2013, which relates urbanity (density measures and land use mix) to centrality values and center potential inherent to the street network. Higher match rate can be interpreted as beneficial.	Y. et al. 2013; Al Sayed et al, 2016
18	Distance matrix	Elasticity of activity levels	Score	Change in activity levels (car counts as proxy) in area i / change in GFA in area i	The indicators make a statement on how activity levels react to changes in GFA or put differently how long it takes for (different) newly developed areas to be used and social integrated in the city. Time lags should be considered in this analysis Values are aggregated to a grid and smoothed within radius of 800 meters, Car counts are taken as proxy for activity levels.	-

Table 10: Indicators of the section *Coherence of plan and market*. Source: Author

Moreover, the correlation between change in GFA and GDP / population/ activity levels is considered and can be interpreted as a proxy for the balance between supply and demand.

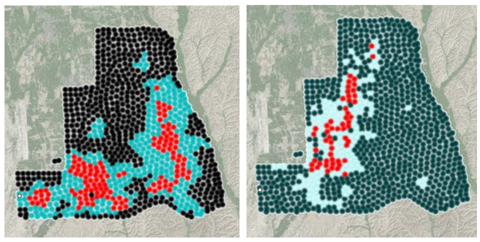


Fig. 57: Areas with high levels of *urbanity* (left) and network integration (right). The indicator *match rate* is calculated using these two data points. From red to dark blue: High to low values, defined by a natural break clustering algorithm. Source: Author

Additional data

For the sake of the analysis of the Lanzhou New Area, both parking and driving cars were counted on the satellite imagery (see Fig. 59). The car counts are used as a (very rough) proxy for the actual activity level. As this analysis is conducted ex-post, only available time-series data can be used, which boils down to satellite imagery. Else wise, it could be easily ensured to 1) improve satellite imagery data quality by frequent flyovers at similar times and weekdays 2) by tapping additional data sources, e.g., by crawling websites or via APIs of mapping services. Taking the case of activity levels, Baidu maps could be used as it offers an API to access real-time activity heat-maps maps of people using Baidu maps (see section Long 2017 and Fig. 58).

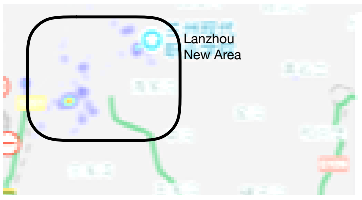


Fig. 58: Real-time activity levels from Baidu maps. Source: Screenshot from Baidu map's mobile app

Fig. 59: Annotated driving cars in the LNA. From left to right: 2010 (160 cars); 2013 (670); 2015 (1260); 2017 (2760). Source: Author

Limitations

Again, it is emphasized that the presented indicators can only be seen as a very first step in the quest to build up a robust evaluation and monitoring system. Ideally, each of them should be validated and computed for many places and projects in order to put returned numbers in graspable relation. Finally, even then, indicators should be interpreted carefully as they depend on the initial data inputs, its quality, and the parameters set by the user.

4.4 Tracing the evolution of the Lanzhou New Area

The spatial development of the LNA unfolded rapidly. Within five years, the naturally grown network of small roads and paths connecting dozens of villages in the Zhongchuan plane were cut through and erased by a strictly geometric system of arterial roads and city highways (Fig. 60) that, in some cases, even exceed whole villages in their width. Road dimension and average block length can be considered as large - even compared to other Chinese new town projects. Figure 62 provides an overview of typical blocks found in the LNA and other cities as well as sections of two road categories (primary and secondary).

Fig. 60: Street network analysis for the Lanzhou New Area. From top to bottom: 2010; 2013; 2015; masterplan. Thickness of the line represents global betweenness centrality; Colors indicate local betweenness (radius = 1800m) ,red indicates high, black low, values. Source: Author

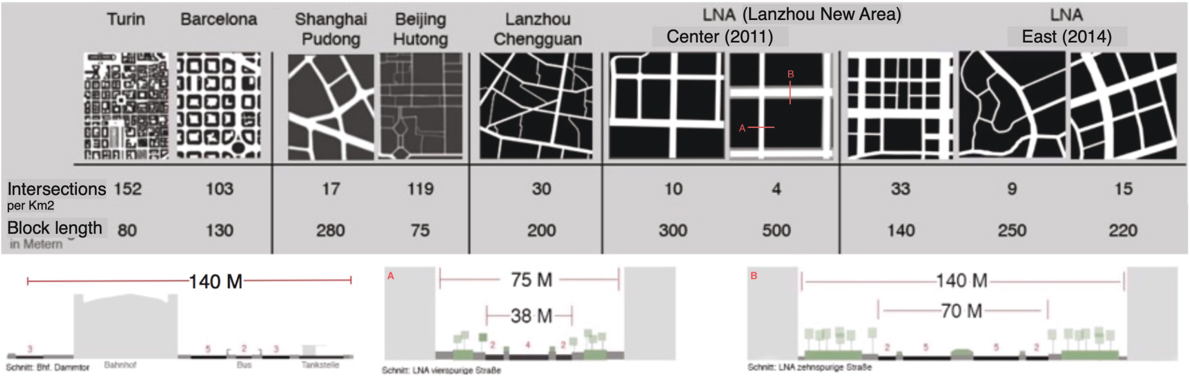


Fig. 61: Sections and Block sizes of different cities and city quarters. Note the differences in the east and west part of the LNA. Section on the left: a typical train station and adherent square in west Germany, Section in the middle and on the right. Two roads in the LNA's center. Source: Author with data from World Bank 2014 and Google.

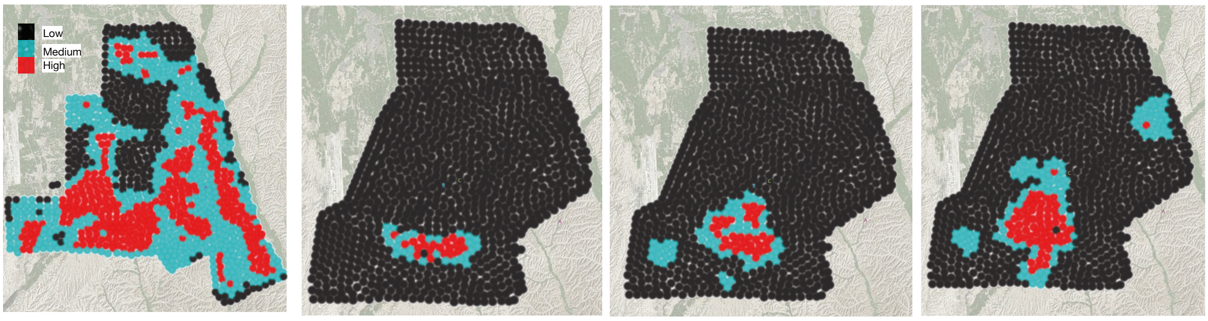


Fig. 62: Visualization of the indicator urban intensity. Values for the masterplan and the years 2013, 2015, 2017 are displayed. Source: Author

The development of the LNA started roughly from the area where network centrality for both the current state as well as the future road network configuration peaks (see Fig. 60). It continues by following the developed road network (and along high centrality values) in northern directions. However, as the density maps in Fig. 64 illustrate, this process happens comparatively un-concentrated, stretching over a radius of 20 kilometers.

From the beginning of 2015, the western part of the LNA starts taking shape. In the north, a large district focused on educational land uses such as universities and vocational training schools, is under construction (Fig. 60). Both in terms of distance to previously developed residential and service-oriented built-up land as well as considering the street network centrality, this newly added part is poorly integrated. It almost appears as the opening of a second potential center for the city. A claim supported by the indicator of *urban intensity* (see Fig. 62). Given the high ratio of yet empty plots in the central areas (not to mention the even lower actual population numbers), this circumstance seems to be sub-optimal. The issue is further illustrated by the indicator *utilized job accessibility* that *scored* between *0.32 and 0.6* (for the period 2013 - 2017), indicating that a denser development performed around two to three times better. The scattered developments caused lost potential.

Fig. 63: Network centrality values of the masterplan. Thickness: betweenness centrality (global); colors: betweenness centrality r = 5000 meter; red indicate high values. Source: Author

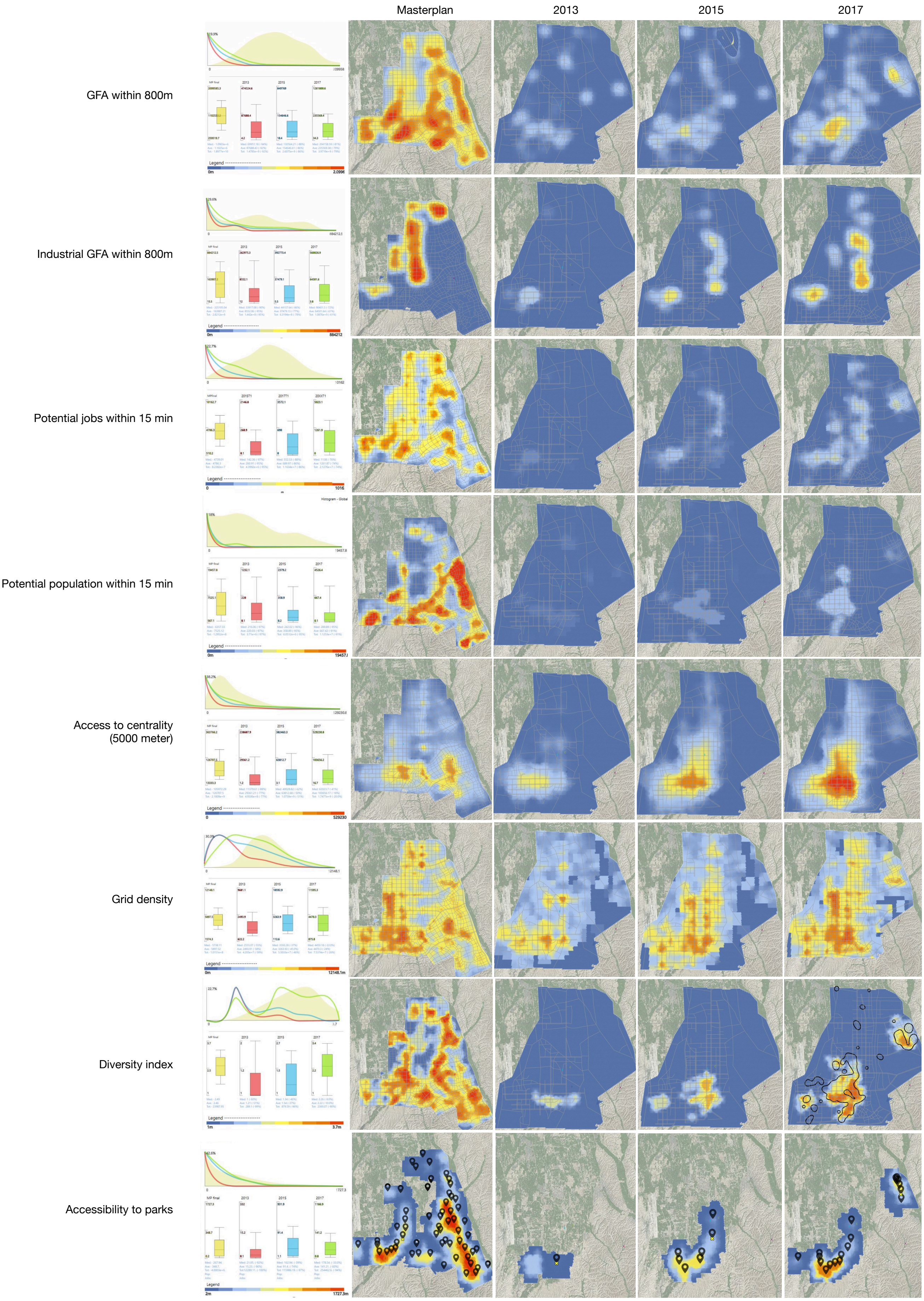


Fig. 64: Visualization various indicators for four development stages. Values for three time steps as well as the situation envisioned by the masterplan are plotted. The left hand side histograms and box plots describe the value distribution of each indicator. Colors represent the time step: Yellow: masterplan; red: 2013; blue: 2015; green: 2017. Visualizations in row five (access to centrality) were not normalized across the four scenarios, thus color ranges varies for each time step. Source: Author

Regarding network centrality in the final configuration (Fig. 63), a general concentration of developments along the west-east axis would have perhaps made more sense. The higher pedestrian and car frequency potentials and intra-accessibility associated with highly integrated areas would have been captured more profoundly. However, it is noteworthy that the planning of the east side was conducted two years after constructions on the west side started. From this, the questions arise if it would have made sense to plan the east-part in stronger adoption to the status-quo of the already developed areas. A crucial aspect in evaluating such questions is the (expected) rate of population and economic growth: In case the town fully develops and populates within, say ten years, the efficiency loss is somewhat diminishing. In contrast, if it takes the city 30 or 50 years to reach its target size, these questions gain significance. The next section taps a little further into this issue by exemplarily comparing an alternative growth trajectory with the observed one. Moreover, the indicators briefly mentioned in this section will be discussed in more detail and comparatively to the alternative scenario.

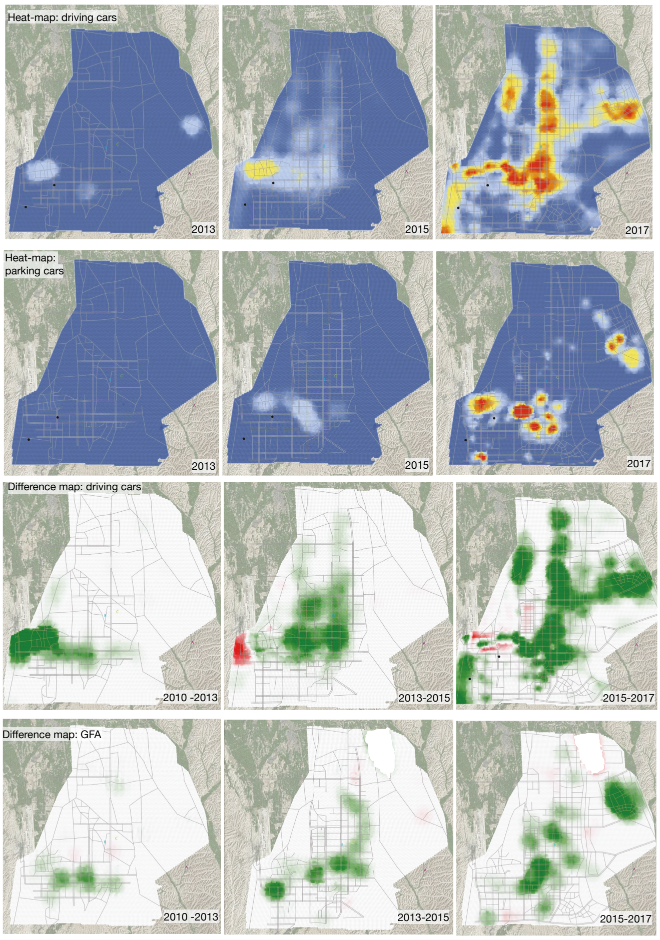
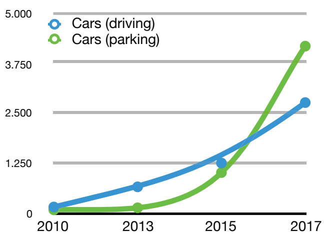


Fig. 65: Changes in car and GFA density. Top-left row: Difference map of counted driving cars. Top-right: Heat-map of driving cars counted. Bottom-left: Difference map of GFA. Bottom-right: Heat-map of parking cars. All visualizations represent densities measured for a radius of 800 meters. Source: Author

So far, the analysis exclusively considered the physical properties of the LNA's evolution. Now, a tentative step toward incorporating a social dimension to the toolboxes analysis is taken. This is approached by relating changes in activity levels (car counts are used as proxy data) with changes in the gross floor area between one period to another (see Fig. 65 and table 11). An indicator to describe the relation could be called Gross Floor Area (GFA) elasticity of activity levels - that is, the relative response of activity levels to a change in GFA within an area. Essentially, this indicator measures how many people are attracted by new

	2013	2015	2017	MP
Population (official estimate) in thousands	120	130	142	1000
avg. potential pop. density	220	360	650	7500
avg. potential job. density	58	280	730	4500
Cars (driving)	670	1260	2760	-
Cars (parking)	150	1024	4160	-



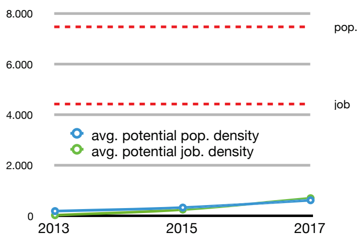


Table 11: Development indicators over time. Densities are calculated as the reachable population/jobs within a 10 min. walk. The red dotted lines depict the densities assumed by the masterplan (MP). Population numbers are from Song 2018. Source: Author with data from Song 2018.

developments and how fast these developments are adopted by users and thus socially integrated into the city.

For the calculation of the indicator, the GFA density (radius = 800) of several land use categories were calculated for each cell of the distance matrix grid (1000). The results are summarized in table 12. Overall, an increase of 10% of GFA in an area is associated with an increase in activity levels between 3,1% and 2%. For educational land uses, this number is significantly higher (17 - 19%). An explanation might be that the government has higher capabilities in allocating jobs and residents (students) to these areas directly. Also noteworthy is the difference between industrial and residential land uses. This perhaps reflects the circumstance that 1) companies reacted to subsidies given for their relocation and 2) that supply and demand for factories usually steer from the same hand in contrast to real estate development were developers and users are two different entities.

Moving one step ahead, the interesting question arises if any spatial properties systematically alter the elasticity of activity levels toward land uses. Robust answers to this question could contribute significantly to the planning of new cities and quarters.

For a first test, the values on activity levels (represented by the number of parking cars) were correlated with street network centrality (both normalized to 0-1000) (see scatterplots in Fig. 66). For both overall GFA and population potential, a positive correlation was found with an r-square of 0.1 and 0.12, respectively. A relation almost absent in the case of educational land uses. Which again would support the point made above, that educational land uses depend less of preferred locations to attract users. Finally, the correlation was repeated using the distance to the LNA center instead of network centrality. Here, the data on the activity level did not show any significant response.

These initial results should be interpreted very cautiously as the underlying data quality as well as quantity (only three periods were looked at) is far from adequate to generalize results. However, further research in this sphere might be of interest and of usability to understand and model urban growth.

Land use category	2013 -2015	2015 - 2017
GFA	0,31	0,2
Residential GFA	0,6	0,5
Industrial GFA	0,8	1,0
Educational GFA	1,9	1,7
Population potential	0,4	0,25
Job potential	0,8	0,8

Table 12: Activity level elasticity for different kind of land uses. A value of 0.31 means that a change in the GFA of the associated land use leads to an increase In activity levels of 0.31% Source: Author

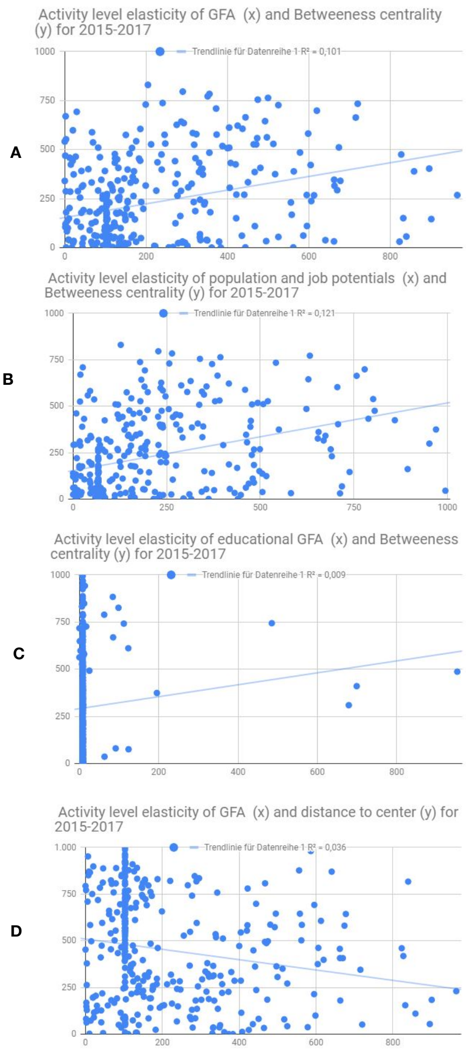


Fig. 66: Correlating activity level elasticity with spatial properties. Activity level elasticities and street network integration (A-C) and distance to the town center (D). The data suggests a slightly positive relationship between activity level elasticity and network centrality. The strength of this relationship varies across different land uses. Values for both activity levels and changes in land uses were normalized between 0 and 1000. Source: Author

4.5 Testing alternative growth trajectories

The third analysis direction further explores the matter of cities' growth trajectories and whether or not there are differences in efficiency and resilience of development paths. Here, resilience refers to the town's performance (as measured by the proposed indicator system) on the way reaching its "final" stage of development the masterplan (and infrastructure) was optimized for.

The analysis will exemplarily be operationalized by comparing the observed growth trajectory of the LNA with an alternative, a simulated one, throughout three development stages. The original LNA masterplan serves as the basis for the alternative scenario as well. For simplicity, the road network is directly taken from the masterplan instead of also being incremental developed. Thus the only difference among the two scenarios is an altered sequence of blocks developed. For each time step, the total share / GFA of land uses was aimed to be as equal as possible across the scenarios to isolate the effect of development sequences on the indicator performance. An overview of job and population densities, as well as the gross floor area density for both scenarios (all measured for a radius of 800 meters), is presented in Fig. 68.

Scenario 1 (S1): In the first two development steps (below named as 2015, 2017) scenario 1 equals the observed development of the LNA. The third development step (here referred to as 2020) is fictional and continues developments around existing hot spots (see Fig. 67).

Fig. 67: Developed blocks and land uses over time. Each dot represents a developed block, its color the associated land use. Yellow: residential; orange: mixed use; purple: logistics/light industry; brown: industry; green: parks; red: commercial. Source: Author

Scenario 2 (S2): This scenario represents a fictional alternative growth trajectory of the LNA. It resembles a compact development, starting from the central business district (CBD). While residential areas are strictly co-developed with near-by commercial and educational facilities, the separation to industrial areas is more significant compared to S1.

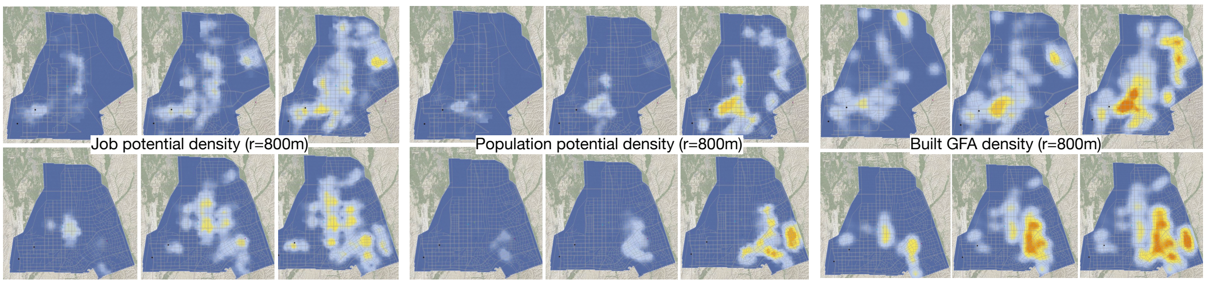


Fig. 68: Density of potential jobs, residents and built GFA over the three periods for both scenarios. Top row: scenario S1; bottom row: scenario S2. Source: Author

Comparing growth trajectories

Both scenarios were modeled and processed in the sketch-mode and then analyzed with the toolboxes evaluation-mode.

One of the most crucial aspect to consider while interpreting the indicators and evaluation options is time and the (assumed) growth rates in terms of population, jobs, etc. In case the city grows exceptionally fast, interim development steps are naturally less significant for welfare and performance estimations. Fig. 69 illustrates this matter by taking the indicator *agglomeration potential* as an example: Both charts show the values for the same two spatial development strategies over a hypothetical period of 16 years. In the right chart, the *final* stage is reached much faster. The benefit of adopting scenario two averages at 10% compared to 30% in a case of slower population growth rates.

As predictions on growth rates are hard to make, especially for new towns, multiple possibilities could be tested. For more risky projects, growth scenarios that perform better in case the city grows slower or remains smaller than hoped for should be considered with greater emphasis. Below indicators of each category will exemplarily be discussed.

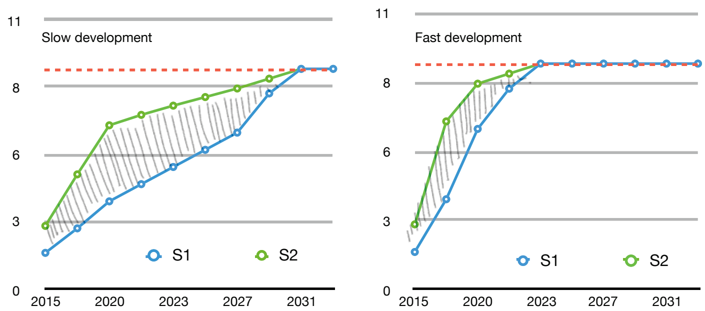


Fig. 69: Two growth paths plotted against different development speeds. Both charts show a hypothetical time series of the indicator agglomeration potential for two different growth paths (S1 and S2). In the left hand side chart a slower development speed of the city is anticipated The surface of hatched area indicates the total benefit of S2 compared to S1. Source: Author

Infrastructure efficiency

In terms of access to amenities, the more densely developing scenario two outperforms S1 in all indicators considered (Table 13). With a strong fluctuation between each period (S2 performs between 54% and 20% better (average of all indicators). The indicators strongly depend on the distribution of residents and the location of amenities (here, schools and parks). The tool and this indicator can, therefore, be employed for optimizing location choices for amenities considering different possible growth paths.

		2015	2017	2020	MP
Schools, closest facility (transit)	S1	46	23,3	23,1	13,4
	S2	29,4	19,2	20,4	
Avg. accessibility to parks	S1	8,8	14	16	30
	S2	9,3	17,7	23	
Share of population with good access to parks*	S1	5	8	8	15
	S2	10	9	12	
Avg. difference S1 to S2		**54,0 %**	**20,1 %**	**35,7 %**	
Categories won:	S1	0	0	0	
	S2	3	3	3	

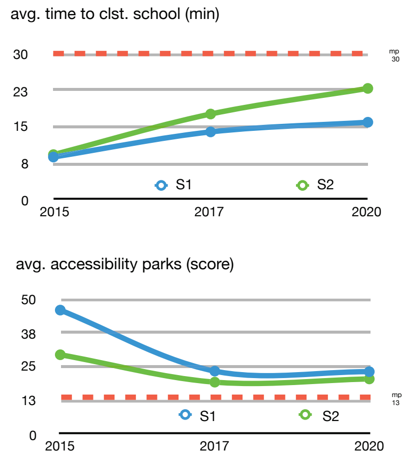


Table 13: Infrastructure efficiencies and growth trajectories. The values of both scenarios are presented for each time step. The color indicates which scenario performed better; green for S2 and blue for S1. Good access to parks is calculated as the distributed population to parks by the gravity function. Source: Author

Social/ urbanity

The results in this category are slightly less clear (Table 14). While not surprisingly, the denser scenario two performs better in terms of land use diversity (Fig. 70). It scores significantly worse in access to centrality for the development stages two and three. As mentioned in the literature review, it is generally beneficial to follow the potentials inherent to the spatial configuration of road networks (e.g., see Karimi 2014). Additionally, the results of the previous section support that argument by the indicated higher elasticity of activity levels to developments near areas with network centralities. Positively, on the other hand, are the better scores of S2 in the first development stage - which could be regarded as especially important for cases such as the LNA where development does not pick up, and it is highly probable for the city to remain in early stages for a longer time.

		2015	2017	2020	MP
avg. agglomeration potential*	S1	1,5	3,6	6,4	16,9
	S2	2,6	6,7	8,2	
Avg. commuting time to jobs (car)	S1	**13,3**	**14**	**14**	15,5
	S2	11	11,8	12,7	
Avg. commuting time to jobs (transit)	S1	53,8	51	55,7	60
	S2	53	43	49	
Ratio	S1	0,24	0,27	0,25	0,26
	S2	0,2	0,27	0,26	
Avg. job accessibility (transit)	S1	1,5	3,62	6,33	16,4
	S2	2,19	6,61	7,7	
Avg. job accessibility (car)	S1	7,65	16,35	28,5	72
	S2	9,3	24,4	30,3	
Ratio	S1	0,22	0,22	0,23	0,22
	S2	0,23	0,27	0,25	
Avg. difference S1 to S2		**28,0 %**	**46,3 %**	**14,8 %**	-
Categories won	S1	1	1	0	
	S2	6	7	7	

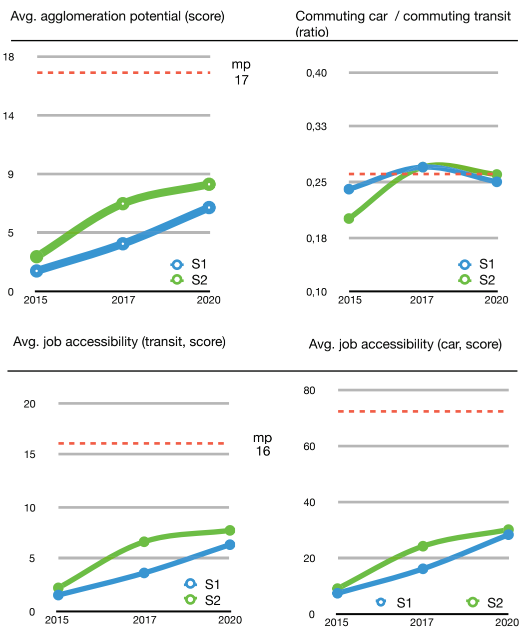


Table 15: Economic indicators and growth trajectories. The values of both scenarios are presented for each time step. The color indicates which scenario performed better; green for S2 and blue for S1. Some indicators are mapped to the charts on the right hand side. Source: Author

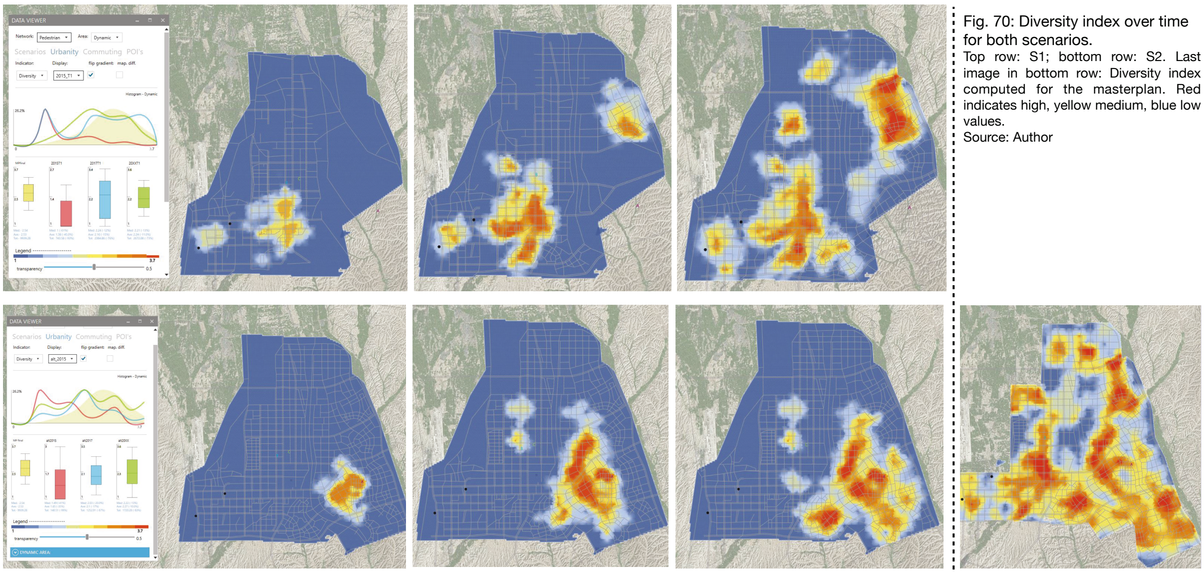


Fig. 70: Diversity index over time for both scenarios.
Top row: S1; bottom row: S2. Last image in bottom row: Diversity index computed for the masterplan. Red indicates high, yellow medium, blue low values.
Source: Author

For the indicator urban intensity and match-rate, again ‚S1 outperforms S2, which is somewhat counter-intuitive: A reason might be behind the way the indicator is computed (see Ye et al. 2013a). By using the natural break method, the indicator perhaps says more about the distribution of values than the actual urban intensity. In the case of S2, the low numbers for medium cells could likely be explained by a more extreme distribution due to denser development (see Fig. 71). Similarly, the match rate, as it is computed using the indicator for urban intensity, should be interpreted cautiously. For more robust results and comparisons across scenarios and cases, it would make sense to replace the natural break method with a scoring system based on the evaluation of multiple cities and defined thresholds for low medium and high levels of urban intensity.

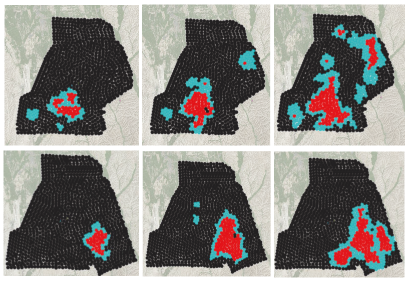


Fig. 71: Urban intensity levels over time.
Top: S1; bottom: S2. Red: high; blue: medium; black: low.
Source: Author

Economy

First of all, the denser and highly mixed-use scenario two overall outperformance the more scattered development patterns of S1. In the critical indicator, agglomeration potential it outperforms S1 nearly by factor two during the first two stages. Afterward, values for both scenarios seem to converge (see Fig. 72). Despite the overall higher remoteness of housing to industrial areas (that play a significant role in the LNA), values for commuting are lower for S2. This can arguably be attributed to the stricter co-development of housing around commercial and office areas. This fact also leads to good performance in avg. job accessibility (transit), especially for the first two stages. The corresponding histogram in Fig. 73 emphasizes the significantly more even (S2) distribution of job accessibility in stage one (represented by the red line). Across all included indicators, S2 performs between 28% and 15% better than scenario one (Table 15).

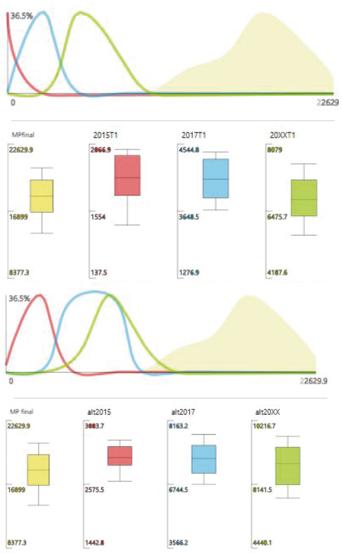


Fig. 72: agglomeration potential over time.
Top: S1; bottom: S2. Source: Author

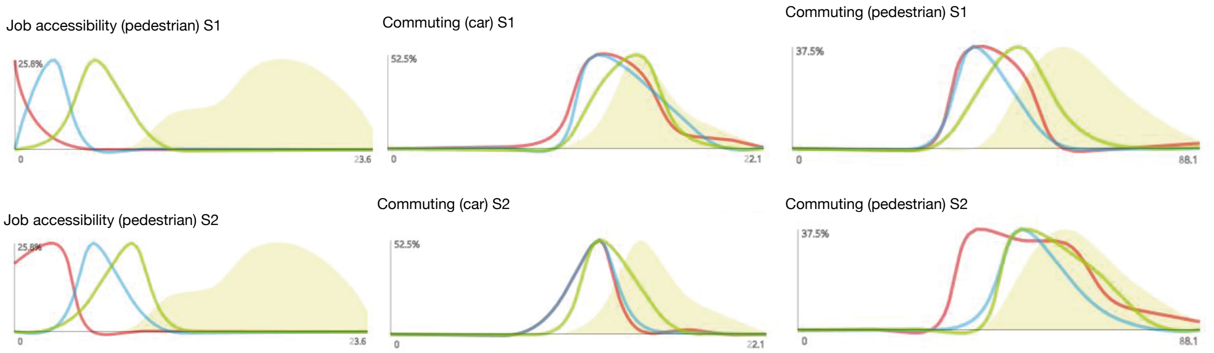


Fig. 73: Histograms of various indicators. Colored lines represent the time steps. Red: 2015; blue: 2017; green 2020; yellow: masterplan. Source: Author

		2015			2017			2020			MP
Avg. diversity*	S1	1,3			2,0			2,2			2,53
	S2	1,6			2,1			2,27			
Avg. access to centrality* (800m)	S1	5,3			10			8			6,7
	S2	5,5			4,7			5,28			
Avg. access to centrality* (5000m)	S1	160			312			266			177
	S2	274			270			251			
Difference S1 to S2:		**18,8 %**			**-23,7 %**			**-14,2 %**			
urban intensity (L/M/H)	S1	90	7	3	83	10	7	65	23	12	26/44/30
	S2	94	3	3	86	6	8	76	12	12	
Matchrate	S1	0.27			0.38			0.61			3.0
	S2	0.12			0.34			0.41			

Table 14: Social/urbanity indicators and growth trajectories. The values of both scenarios are presented for each time step. The color indicates which scenario performed better; green for S2 and blue for S1. Source: Author

In conclusion, the analysis suggests that scenario S2 is preferable over S1, while simultaneously weaknesses of S2 were uncovered. Such as the low access to network centrality in development stages two and three. These results could be seen as a starting point for further and systematic exploration of alternative growth trajectories and their likely performance and for the question in which direction to continue from the current spatial structure and growth prospects of the Lanzhou New Area.

Reflection

This section gave an exemplary outlook on how the toolbox can be applied in evaluating, comparing, and designing different trajectories for city's spatial development, thus, facilitating more resilient growth processes. A strength of the suggested analysis is its integrated view on top-down (Masterplan) and bottom-up (where are new developments likely to happen? To which areas would people move to first?) development processes and their interaction effects. Moreover, a time dimension in evaluating the impact and risk of infrastructure projects such as highways or the location of schools can explicitly be considered.

For future iterations, the analysis scope of the toolbox could be extended by adding simulation modules, such as for land use and retail distribution. On the one hand, this could help (if the model's parameters are calibrated meaningfully) to expose natural growth paths (e.g., in relation to the region's topography, road network, nearby settlements, etc.). On the other hand, the interaction with actual population counts could be better integrated: by balancing the actual population and the demand, they pose toward other land uses and their spatial allocation. Again, as the number of parameters rises, results would need to be interpreted carefully as trends and probabilities.

5. Conclusion

Following the call for a more evidence-based approach (Karimi 2014) on urban planning that pays more attention to the self-organizational forces inherent to cities (Bertaud 2018), a prototype of an integrated toolbox for spatial analysis, scenario sketching, and monitoring were developed and exemplarily demonstrated across three use-cases of Lanzhou New Area.

The framework is positioned to assist urban design and planning efforts in cases the employment of complex models is not possible due to budget, data, or know-how constraints. The implementation in the Rhino/Grasshopper environment allows for interactive, indicator guided design processes and analysis. Besides, the modular framework can easily be expanded and integrated with other models created in Grasshopper.

The backbone of the framework is a weighted graph-gravity model, allowing to simulate interaction effects between land uses associated socio-demographic properties, and points of interest through a multi-modal transportation network.

The toolbox can be applied in various use-cases (Fig. 74). Three examples were given: Starting with reasoning a preferable location for a second railway overpass in Kagan, the framework was employed to conduct an impact analysis of infrastructure projects.

Secondly, the Lanzhou New Area's evolution was traced and quantified across four periods. Finally, taking the observed growth trajectory as a comparison, an alternative spatial growth scenario for the LNA was simulated, and suggestions were laidout in which direction to facilitate spatial growth in order to increase resilience toward unexpected development trajectories.

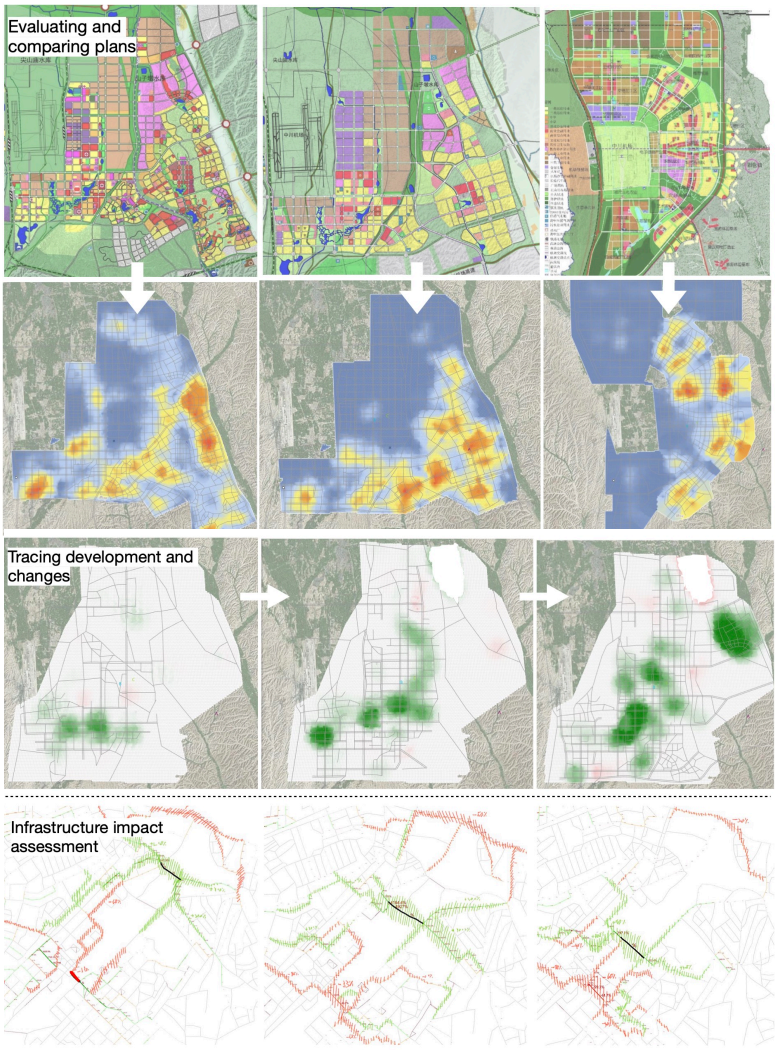


Fig. 74: Application possibilities of the toolbox. Left: Indicator driven comparison of the three masterplan proposals; Top-right: Monitoring and visualizing changes in GFA; Bottom-right: Assessing the impact of several bridge locations on pedestrian flow potentials. Source: Author

Regarding the research questions formulated in the introduction the results can be summarised as follows.

Q1 - Do spatiotemporal configurations matter in regard to the efficiency of urban form measured over time?

Analysis of the hypothetical scenario delivered clear hints to a high level of inefficiency inherent to the observed, large scaled and scattered spatial development of the LNA (in regard to the proposed indicator system).
The hypothetical scenario scored quite differently and especially superior (18 to 54% better scores across measured metrics) in early development stages of the town (with a small population and job numbers). When looking at the performance of the city as envisioned by the masterplan, it scored best in a range of indicators, for instance land use diversity or access to parks and schools, indicating that the masterplan indeed works best at its final state.
As both scenarios were grown to the targeted size of one million habitants, the performances logically converged. However, the alternative scenario did so with a better performance when aggregated over time. Especially for cases of slow development, as observed in the case of the LNA, such difference matters and should be more deeply considered in planning processes early on to grow cities that work well during all stages of development.

Due to the limited scope of this thesis, the observed spatial development was compared to only one alternative scenario. Ideally, multiple scenarios should be tested and evaluated to iteratively get closer to a near-optimal design solution that can serve as a benchmark.
Moreover, presented and quantified inefficiencies of the LNA's observed spatial development trajectory should be interpreted carefully and in relation to other spatial configurations or case-studies (which do not exists yet, see below). Here, only the phasing of developed plots was altered between both scenarios, for more complex analysis other spatial features such as the street network, land uses and densities should be subject to adjustment as well. For more limitations, resulting from the early development stage of the toolbox and data availability see below.

Q2 - Which rules can be deduced to guide the design of spatiotemporal configurations that perform efficiently throughout all stages of the urban growth process?

The analysis conducted in chapter 4 suggests that, indeed, more efficient (in regard to the indicator system) spatiotemporal development paths can be envisioned compared to the observed one. Moreover, the data provide hints to the existence of potentially generalizable rule-sets informing strategies for spatial growth trajectories that are more resilient, given uncertain growth projections. Hereof, centrality measures, calculated with the weighted, multimodal graph-model appear as especially promising: According to observations made in chapter four, a new town construction should start from the most integrated and central areas (measured by local radii) in the transportation network and expand along with high centrality measures as the radius of analysis is gradually expanded. Moreover, a first glance into the elasticity of activity levels supported this position as plots associated with high centrality values appeared to attract human activity more quickly. It should be emphasized that these results, as of now, are highly hypothetical and need further investigation with more cases and data to be verified.

Besides, many more factors that have been omitted so far should be studied in-depth, such as: Experimenting with different transportation networks and land use configurations when designing alternative scenarios. On the one hand, to optimize existing designs and on the other hand, to search for generalizable rules or approaches for designing cities based on these parameters.

In case an efficient strategy is found, the issue remains how to facilitate developments accordingly. One way to proceed is allocating public buildings correspondingly. For instance, this might include to give up the concept of a large university campus or town in the outskirts (as seen in the LNA) in favor of multiple small campuses in the city center. Certainly, there are tradeoffs to make that partially depend on a city's development speed (see chapter 4.5) and, among other, traffic-related issues. Another way is the definition of obligatory planing codes - or, preferentially, the already mentioned second approach to planning through incentives (e.g. monetary) and by anticipating location-choice behavior by stakeholders such as property developers or retailers based on the spatial configuration of the city. For example, by indirectly planning the location of retail centers through means of the network configuration and density distribution instead of land use plans. These are approaches, the suggested toolbox provides the basic functionality for.

Discussion

The toolbox made a first step to integrate designing, plan evaluation, and monitoring of actual urban development (the bottom-up response to making plans so to say) into one, unified analysis framework.

On the one hand, evaluating plans employing this kind of integrated analysis and simulation significantly contribute to efficiency and the resilience of urban development dynamics. On the other hand. On the other hand, constant tracking of a city's development enables to move toward a paradigm of just-in-time planning, where decisions (e.g. the location of public amenities) are made when needed, which arguably minimizes the need of making long term forecasts and related uncertainty - in the way of the second approach to urban planning mentioned in the introduction.

It is important to emphasize that the toolbox at no point claims to cover all issues related to and essential for planning and monitoring urban systems. Thus, it can only contribute to the fields it makes dedicated statements on. Moreover, its accuracy crucially depends on the data quality used as input and the calibration of its parameters. As both can be challenging, the toolbox can be seen as a planning and design decision support system - as one additional tool to base choices on. Similarly, as of now, results from case-study should be seen in the light of serving to exemplify and to explore application possibilities of the toolbox. As the observed timeframe of five years is short, and only one city was investigated, the study's findings are not generalizable yet but should rather be seen as a departure point for further research.

All in all, the integration of design, evaluation, and monitoring into one framework appears as a promising approach toward more evidence-based planning. The implementation in Rhino/Grasshopper enables easy expandability with additional task-specific analysis modules or the quick set-up of interactive environments for co-designing - for instance, in augmented reality (see fig. 34 on page 39).

Moreover, there are several directions how the toolbox could be extended and improved:

- Implementation of modules for basic land use development simulations and retail allocation
- A module for basic traffic estimations (see Fig. 19)
- A module to integrate the regional context into the analysis of cities (see Fig. 75)
- Building up a catalogue of indicators and their implementation
- An incorporation of big-data sources such as from social media or online mapping services and deep-learning based automatized feature extraction from high-resolution satellite imagery (volumes of buildings, land uses, number of cars etc., see fig. 45)
- A web-based evaluation-mode interface for rapid sharing of results and more collaborative workflows

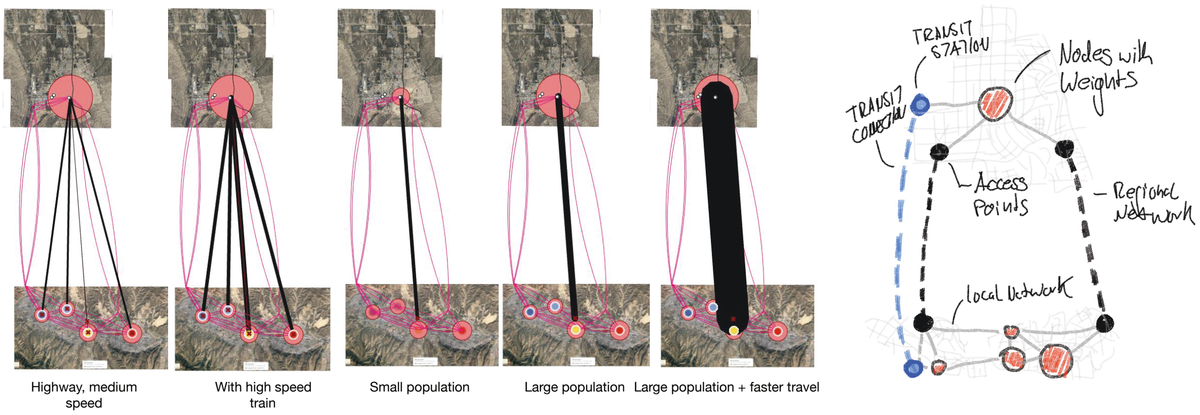


Fig. 75: Incorporating the regional scale (concept). Left: Impact of different parameters on the interaction between two regions. The line's thickness indicates the level of interaction (eg. number of commuters). Right: A simplified graph for the regional scale is created. Three kinds of nodes exist: Nodes that carry weights (such as number of jobs and population) which act as attractors (depicted in red), and two kinds of access points for transit and vehicles (blue, black). Exchange rates between regions are calculated on the simplified network. In access points, currents (e.g. commuters) in both directions are tracked and forwarded as weights to nearest nodes on the detailed road network. Source: Author

Several steps should be taken to validate the proposed framework, for instance, by employing the toolbox systematically in the analysis of several cities and through cross-checks against results from similar analysis, models, or data.

Finally, departing from the findings of section 4.5, further elaboration is needed which, and to what degree, spatial features mediate the pace, new developments are socially integrated into a city or initiate being used and populated. This could be looked at for both local (neighborhood) and regional (cities) areas and might be of interest, especially for the development of new towns.

6. List of Tables and Figures

Tables

Figures

7. References

Achary A., Karimi K., Parham E., Guven A., Uyar G. (2017): City planning using integrated urban modeling. Jeddah Structure Plan. Proceedings of the 11th Space Syntax symposium

Al Sayed K., Penn A. (2016): On the nature of urban dependencies: How Manhattan and Barcelona reinforced a natural organisation despite planning intentionality. Environment and Planning B, 43

Batty M. (2008): Fifty years of urban modeling: Macro-statics to micro-dynamics. In: Albeverio S., Andrey D., Giordano P., Vancheri A. (eds.).The dynamics of complex urban systems. Physica-Verlag, Heidelberg

Batty M. (2017): The new science of cities. MIT Press, Cambridge

Bertaud A. (2018): Order without design. How markets shape cities. MIT press, Cambridge

Bielik M., Schneider S., Kuliga S., Griego D., Ojha V., König R., Schmitt G., Donath D. (2018): Measuring the impact of street network configuration on the accessibility to people and walking attractors. Networks and Spatial Economics, 18, 657-676

Blake P. (1977): Form follows fiasco: Why modernist architecture hasn't worked. Little Brown, Boston

China Academy of Urban Planning (CAUPD) (2015): 兰州新区总体规划(2011-2030) 实施评估报告. www.docin.com/p-1124452318.html, accessed: 25.1.2019

China National Bureau of Statistics (NBS) (2018): Shanghai Pudong new area GDP. CEIC Data. https://www.ceicdata.com/en/china/gross-domestic-product-municipality-district/cn-gdp-shanghai-pudong-new, accessed: 5.5.2019

China-up (2013): 兰州新区总体规划. http://www.china-up.com/podcast/archives/866, accessed: 05.06.2018

Ding K, (2019): 兰州新区晒"成绩单"：增速排名19个国家级新区第一. http://www.sohu.com/a/294965759_123753, accessed: 25.04.2019

Dogan T., Samaranayake S., Saraf N. (2018): Urbano: A new tool to promote mobility-aware urban design, active transportation modeling and access analysis for amenities and public transport. SimAUD proceedings

Ewing R., Cervero R. (2010): Travel and the built environment. Journal of the American Planning Association, 76, 265 - 294

Florida, R. (2004): Cities and the creative Class. Routledge, New York

Gehl J. (2010): Cities for people. ISLAND PR, Washington D.C.

Geurs K.T., Van Wee B. (2004): Accessibility evaluation of land-use and transport strategies review and research directions. Journal of Transport Geography, 12(2), 127-14

Graham D.J., Gibbons S., Martin R. (2010): The spatial decay of agglomeration economies: estimates for the use in transport appraisal. http://personal.lse.ac.uk/gibbons/papers/agglomerationreport.pdf, accessed: 23.01.2019

Hall P. (1996): Cities of tomorrow: An intellectual history of urban planning and design in the twentieth century. Blackwell Publishers, Oxford

Hansen W.G. (1959): How accessibility shapes land use. Journal of the American Institute of Planners, 25(2), 73 - 76

Hensher D. A. (2004): Handbook of transport geography and spatial systems. Elsevier, Amsterdam, Boston

Herthogs P., Tuncer B., Schläpfer M., He P. (2018): A weighted graph model to estimate peoples presence in public space. eCAADe proceedings

Hillier B. (1996): Space is the machine: a configurational theory of architecture. Cambridge University Press, Cambridge

Hillier B. (2009): Spatial sustainability in cities. Organic patterns and sustainable forms. Proceedings of the 10th Space Syntax Symposium

Hillier B., Penn A., Hanson J., Grajewski T., Xu J. (1993): Natural movement - or configuration and attraction in urban pedestrian movement. Environment and Planning B: Planning and Design 20, 29–66

Hillier B., Julienne H. (1984): The social logic of space. Cambridge University Press. Cambridge

Hospers G. J. (2003): Creative cities: Breeding places in the knowledge economy. Knowledge, Technology & Policy, 16, 143-162

Huff D. (1963): A probabilistic analysis of shopping center trade areas. Land Economics 39, 81-90

Kairanbay M., Jane M. H. (2013): A review and evaluations of shortest path algorithms. International Journal of Scientific & Technology Research 2, 99-104

Karimi K, Vaughan L. (2014): An evidence-based approach to designing new cities: the English New Towns revisited, in: Explorations in urban design: An urban design research primer, 261-26, Ashgate, Surrey

Kroll C. (2008): Brasilia or the limits of theory. Agora Journal of Urban Planning and Design, 23 -27

Lanzhou bureau of planning (LBP)(2015): 兰州新区总体规划(2011-2030) (2014修改). http://www.lzxqgt.gov.cn/lzxqghweb/publish/cxgh/ 20151118095100.html, accessed: 05.06.2018

LNA Administration Committee (LNAAC) (2015): 兰州新区总体规划2014修改简介. http://www.lzxqgt.gov.cn/lzxqghweb/publish/cxgh/20151118095100.html, accessed: 21.01.2019

Long Y., Li M., Li J. (2018): Monitoring built environment of China with new data: Indicator system and case studies. Urban Development Studies 25(4)

Long Y., Mao Q., Shen Z. (2013): Urban form, transportation, energy consumption, and environment impact integrated simulation: A multi-agent model. In: M. Kawakami et al. (eds.), Spatial planning and sustainable development: Approaches for achieving sustainable urban form in Asian cities. Springer Science+Business Media, Dordrecht

Marcus L. (2007): Spatial capital and how to measure it. An outline of an analytical theory of the social performativity of urban form. Proceeding of the 6th Space Syntax Symposium

Melo P, Graham D. J., Noland R. B. (2009): A meta-analysis of urban agglomeration economies. Regional Science and Urban Economics, 39, 332-342

National Development and Reform Commission (NDRC) (2015): 国家级新区发展报告2015. China Planning Press, Beijing

Newman M.E.J. (2010): Networks: An introduction. Oxford University Press, Oxford

Nowaczynski A. (2017): Deep learning for satellite imagery via image segmentation. Deepsense.ai, accessed: 10.2.2019

Offenhuber D., Ratti C. (2015): Decoding the city: Urbanism in the age of Big Data. Birkhauser, Basel

Oishi S., Saeki M., Axt J. (2015): Are people living in walkable areas healthier and more satisfied with life? Applied Psychology: Health and Well-Being, 7(3), 365–386

Oliveira, V. (2016): Urban morphology: An introduction to the study of the physical form of cities. Springer, Heidelberg

Piovani D., Arcaute E., Uchoa G., Wilson A., Batty M. (2018): Measuring accessibility using gravity and radiation models. Royal Society Open Science, 5(9)

Planning bureau of Suzhou (PBS) (2015): Technical standards of transportation impact analysis in Suzhou (Chinese). http://www.szghj.gov.cn/front_jtyxpjzl/secondindexnosub.html, accessed: 19.02.2019

Porta, S., Latora, V., Wang, F., Rueda, S., Strano, E., Scellato, S., Latora, L. (2012): Street centrality and the location of economic activities in Barcelona. Urban Studies, 49(7), 1471–1488

Redmon J., Farhhadi A. (2018): YOLO 3: An Incremental Improvement. https://arxiv.org/abs/1804.02767, accessed: 19.10.2018

Rode P., Floater G. (2014): Accessibility in cities: Transport and urban form. NCE Cities, 3, 239-273

Roland Berger Strategy Consultants (2009). Die Go-West-Strategie der chinesischen Regierung. http:// www.rolandberger.com/media/pdf/Roland_Berger_Go_West_Strategie_Chinas_20100113.pdf, accessed: 12.08.2017

Sevtsuk A., Kalov R. (2017): Patronage of urban commercial clusters: A network based extension of the Huff model for balancing location and size. Environment and Planning: Urban Analytics and City Science, 45(3), 508–528

Shepard W. (2015): Ghost cities of China: The story of cities without people in the world's most populated country. Zed Books, London

Shih L., Rudolf M. (2014): Machtzentralisierung im Eiltempo. China Monitor 13. http://www.- merics.org/merics-analysen/infografikchina-mapping/seidenstrassen-initiative.html, accessed: 27.12.2018

Song S. (2018): 兰州新区之困：人口跟不上规划 600万平米库存需消化. In: Xinlang zonghe. https://finance.sina.-com.cn/china/dfjj/2018-03-30/doc-ifystwmu2382134.shtml, accessed: 25.04.2019

Sprengel E. (2017): Segmenting Aerial Images using Neural Networks. Medium.com, accessed: 10.2.2019

Stanilov, K. (2010): Bridging the gap between urban morphology and urban modeling. Urban Morphology, 14, 123-124

Thomas I., Verhetsel A., Hammadou H., Vanhofstraeten D. (2003): Distance decay in activity chains analysis. A Belgian case study. ERSA conference papers, 43

Tobler W. (1970): A computer movie simulating urban growth in the Detroit region. Economic Geography 46, 234–40

Van Etten A. (2018): Broad area satellite imagery semantic segmentation (BASISS), extracting road masks from massive space net images. Medium.com, accessed: 10.1.2019

Wang F., Anipova A., Porta S. (2011): Street centrality and land use intensity in Baton Rouge, Louisiana. Journal of Transport Geography 19, 285 - 293

Wilson L., Danforth J., Cerozo Davila C., Harvey D.(2019): How to generate a thousand master plans: A framework for computational urban design. SimAud proceedings, 113 - 120

World Bank; Development Research Center of the State Council, the People's Republic of China (2014): Urban China: Toward efficient, inclusive, and sustainable urbanization. World Bank, Washington DC

Wu F. (2015): Planning for growth: Urban and regional panning in China. Routledge, New York

Wu H., Yang Z. (2015): 从国家级新区战略看过架战略空间演进. 城市发展研究 3/ 2015, 1-10

Xu C. (2017): 国家级新区兰州新区发展路径的再思考实践探索. www.sohu.com/a/138487164_611316, accessed: 10.03.2019

Yang T. (2015): A study on spatial structure and functional location choice of the Beijing city in the light of Big Data. Proceedings of the 10th Space Syntax Symposium, London

Yu Y., van Ness A. (2013a): Measuring urban maturation processes in Dutch and Chinese new towns. The Journal of Space Syntax, 4/2013, 18 -37

Yu Y., van Ness A. (2013b): The spatial flaws of new towns: Morphological comparison between a Chinese new and old town through the application of space syntax, spacematrix and mixed use index. Proceedings of the 9th international Space Syntax Symposiums, London

Zhang Y. (2017): City matrix. An urban decision support system augmented by artificial intelligence. Master's thesis. Massachusetts Institute of Technology, Boston

Zuend, D., Woodbury, R., & Schmit, G. (2016). Meso-scale modeling of residential and business locations. *SIMULATION*, *92*(3), 295–306

Affidavit

I hereby affirm that this Master's Thesis represents my own written work and that I have used no sources and aids other than those indicated. All passages quoted from publications or paraphrased from these sources are properly cited and attributed.

The thesis was not submitted in the same or in a substantially similar version, not even partially, to another examination board and was not published elsewhere.

5.6.2019

Date

Signature