EVIDENCE-BASED TRANSIT AND LAND USE SKETCH PLANNING USING INTERACTIVE ACCESSIBILITY METHODS ON COMBINED SCHEDULE AND HEADWAY-BASED NETWORKS

Matthew Wigginton Conway, Corresponding Author
Conveyal
641 S St. NW, Third floor, Washington, DC 20001 USA
Tel: 240-230-6529; Email: mconway@conveyal.com

Andrew Byrd
Conveyal
641 S St. NW, Third floor, Washington, DC 20001 USA
Tel: 240-230-6529; Email: abyrd@conveyal.com

Marco van der Linden
Movares Nederland B.V.
Daalseplein 100
Postbus 2855
3500 GW
Utrecht, Netherlands
Tel: 0613942998; Email: marco.vd.linden@movares.nl

Word count: 6,499 words + 4 figures x 250 words each = 7,499 words

November 15, 2016

ABSTRACT

There is a need for indicators of transportation-land use system quality that are understandable to a wide range of stakeholders, and which can provide immediate feedback on the quality of interactively designed scenarios. Location-based accessibility indicators are promising candidates, but indicator values can vary strongly depending on time of day and transfer wait times. Capturing this variation increases complexity, slowing down calculations.

We present new methods for rapid yet rigorous computation of accessibility metrics, allowing immediate feedback during early-stage transit planning, while being rigorous enough for final analyses. Our approach is statistical, characterizing the uncertainty and variability in accessibility metrics due to differences in departure time and headway-based scenario specification. The analysis is carried out on a detailed multi-modal network model including both public transportation and streets. Land use data are represented at high resolution. These methods have been implemented as open-source software running on commodity cloud infrastructure. Networks are constructed from standard open data sources, and scenarios are built in a map-based web interface.

We conclude with a case study, describing how these methods were applied in a long-term transportation planning process for metropolitan Amsterdam.

Keywords: public transit, accessibility, scenario planning

INTRODUCTION

It is often difficult to understand the impact of hypothetical changes to transportation and land use systems. Planning decisions are frequently made in the absence of clear evidence of the advantages and disadvantages of different scenarios. Traffic models are used to provide feedback on proposed projects, but model turnaround times can be prohibitive for use in interactive sketch planning exercises.

Others have concluded that the most important feature missing from quantitative planning tools is real-time interactive scenario evaluation, with maps and other graphical displays for ease of communication *(1)*. With these findings in mind, we aimed to implement indicators of transportation-land use system quality that are understandable to a wide range of stakeholders, and which provide immediate feedback on the quality of interactively designed scenarios, without compromising measure quality. We have chosen to focus on cumulative opportunities accessibility indicators (e.g. number of jobs reachable in 45 minutes), which we compute near-instantaneously by applying novel optimizations and parallelization to existing routing algorithms.

The underlying model is multi-modal, including a street network for bicycle and pedestrian movement, as well as detailed timetabled public transportation. Analysis is carried out at high spatial resolution to avoid aggregation error. Networks can be prepared quickly in any location with General Transit Feed Specification (GTFS) transit schedules and street data from OpenStreetMap. We apply Monte Carlo methods to confront the inherent uncertainty in travel time and accessibility results that occur when transit routes are specified using headways rather than exact schedules in scenarios due to routes being in or out of phase. These goals are achieved by adopting a statistical approach to travel times and accessibility indicator values, treating them as distributions instead of assigning them a single value.

The whole system appears to its users as a modern web application. Modifications can be made to the network interactively using a map-based scenario creation tool. Several key optimizations make this system fast enough to provide feedback on the regional accessibility impact of a scenario within minutes. It performs calculations in parallel using commodity cloud services. This enables collaborative scenario planning where ideas are interactively sketched up and assembled into proposals based on rapid map-based quantitative feedback. All software implementing these methods is available under a permissive open source license, bringing transparency and reproducibility to planning processes.

This paper will first present these accessibility methods in detail, then describe how they were used in regional planning in the Netherlands.

CHOOSING AN ACCESSIBILITY METRIC

The term “accessibility” is quite generic, simply referring to any measure of an ability to access destinations. Many metrics satisfy this definition *(2, 3)*. We have chosen to use a relatively simple cumulative accessibility metric, counting the number of opportunities within a particular travel time of each origin, with no distance decay function. This metric has been used in the past, notably by the Accessibility Observatory at the University of Minnesota *(4)*.

Owen and Levinson *(4)* observe that when working with transit networks, accessibility can be highly variable over time, especially when frequencies are low. After a vehicle passes, it may be many minutes before another arrives; the accessibility from a location the minute before a vehicle arrives is much higher than the accessibility the minute after it leaves. To avoid making results sensitive to exact departure time, we take their approach of considering accessibility at many departure times within a time window.

Owen and Levinson compute the accessibility at each minute, then take an average. Opportunities reachable on average is a valid metric for determining the usefulness of a transit system when people leave randomly (i.e. based on factors uncorrelated with transit schedules), rather than building their lives around the schedule. One desirable property of the average is that it is sensitive to service frequency. When the frequency of transit vehicles is increased, the average accessibility will increase because wait times decrease. This is desirable because wait time is a major portion of a transit trip, especially for people who wish to use the system spontaneously, without consulting a schedule.

There are two ways to compute the average accessibility. One is to compute the accessibility for a particular origin at every minute over the time window and take an average, as Owen and Levinson do. However, this treats opportunities as being completely fungible, even on a minute-by-minute basis. Consider a location which has infrequent trains in two directions. During the first half of each hour, 100,000 jobs in a city to the south can be reached within the travel time budget, and during the second half of each hour, 100,000 jobs in a city to the north can be reached. If accessibility is computed at each minute and averaged, the result will be that 100,000 jobs are accessible at the average departure minute. While strictly true given the definitions above, this implies that people may choose their job based on what time they happen to leave home.

Therefore, we have chosen to instead compute accessibility using average travel times. We calculate the travel time to each destination at each minute, take an average of those travel times at each destination, and sum opportunities in destination cells that have an average travel time less than the specified time budget.

However, the arithmetic mean poorly characterizes destinations that are completely inaccessible for part of the time window. Initially, we averaged only the minutes in which the destination was accessible. This can yield unreasonable accessibility results when there is transit service for only a small part of the time window. For example, a location that is served by a single direct bus 5 minutes after the start of the time window will show extremely high accessibility, because the wait time for the vehicle is extremely short *during those minutes of the time window in which the vehicle is usable at all*, and other minutes are excluded from the calculation. Excluding these minutes biases the results by ignoring what would otherwise be very long travel times.

This problem is apparent when comparing scenarios. Suppose that all-day local service was added alongside that single direct bus. If the local service was slower than the express, accessibility would decrease because higher travel times would be included in the mean for previously excluded departure minutes. This is clearly incorrect; adding service should never cause a decrease in accessibility. Indeed, this is one of Geurs's criteria for an accessibility metric *(2)*.

One solution is to always consider all locations reachable, and allow walking 20 kilometers or waiting until the next day to catch a bus. However, these are trips very few would consider taking, so including them in average travel times seems incorrect. Another option is to specify a percentage of the time window during which a destination must be accessible to be included in averages at all. However, results can be quite sensitive to this parameter, and with low values increases in service can still cause decreases in accessibility.

Our preference is to use median travel time, rather than mean. Travel times for departure minutes at which the destination is unreachable can then be considered infinite and sorted above the longest travel times. To compute the number of destinations reachable within a given median travel time, one need only know whether the travel time to each destination at each departure minute is greater than or less than the travel time cutoff; when a destination is unreachable its travel time is simply considered to be above the cutoff.

Using the median also alleviates concerns about maximum trip length. When computing the mean, trips above a certain length are considered impractical and eliminated. As outliers, these very long trips have a strong influence on the mean, and changing the maximum trip length parameter can have a drastic effect on mean-derived results. Defining our indicator in terms of a median travel time threshold alleviates this concern because all travel times greater than the threshold (whether outliers, trips exceeding the maximum length, or impossible trips) influence the result only by their position above the threshold, not by their specific magnitude.

CREATING SCENARIOS

In our view, accessibility measures are most useful for evaluating proposed changes to the land-use/transportation system. For example, one might want to test the impacts of a new rapid transit line, or of increased frequencies on an existing bus system. A multitude of studies have used accessibility to evaluate scenarios *(5, 6, 7)*.

One challenge in these analyses is creating the scenarios themselves. Our tools use GTFS data as input, but GTFS can be difficult to create *(7)*. GTFS is geared towards transit operations and contains much more data than is typically available during the planning process, for example exact schedules, stop names, etc. It can take several minutes to generate a routable network model from GTFS, which is not tenable for rapid-turnaround sketch planning where the effects of a change to the network should ideally be visible within seconds of beginning to draw it.

Thus, we exploited the fact that most changes are small relative to the size of the full network, and created a lightweight vocabulary for expressing these modifications. It consists of several fundamental operations: reroute, add a new route, adjust speed, adjust dwell time, adjust headway, remove stops, and remove trips. These modifications can be applied to a routable network almost instantaneously. We have implemented a map-based graphical interface to produce scenarios expressed in this vocabulary. We plan to expand the set of operations to encompass changes to land use and the street network.

SELECTING A PATHFINDING ALGORITHM

Our accessibility indicator requires us to determine travel times from each location to every other location in a region. The path-finding algorithms used to find travel times are the main speed bottleneck of accessibility tools, so much effort went into designing and optimizing this part of the system.

There are two broad categories of techniques for pathfinding in transit networks: schedule-based algorithms and frequency-based algorithms *(8)*. Schedule-based algorithms work on timetables, while frequency-based algorithms work with a simplified representation of the transit network containing only travel times and headways.

There are advantages to working with timetable algorithms. Our source data are GTFS timetables, so in order to use a frequency-based algorithm headways must be inferred for each line. Variation in headway or travel time on a route over the time window is lost, as well as any phase information between lines (for example, buses timed to meet trains). The frequency approach is particularly questionable when there are infrequent services.

Another challenge with using frequencies is the common lines problem *(9)*. If there are *n* lines providing service between one location and another, the expected travel time is not simply half of the headway plus the in-vehicle travel time, but rather the minimum of several probability distributions for the arrival time on each line. Several authors have explored how to address this problem analytically *(9, 10, 11, 12)*.

Some *(10, 11)* propose the use of hyperpaths, combinations of paths laid on top of each other. They point out that the common lines problem is more general than previously described: not only can there be multiple competing lines on a corridor, but there can also be multiple competing paths to a particular destination that may not geographically share any common segments. For example, in a gridded network, to travel northwest, one could either take a northbound line then a westbound line, or a westbound line followed by a northbound line. This makes the problem significantly more complex, and brings up the question of how to group paths into choice sets. While solving it is not impossible, we avoid the issue entirely by working with the original schedules, which makes the choice set all paths that could yield the earliest arrival given a departure at some point in the time window.

We use the RAPTOR algorithm for transit pathfinding *(13)*. This algorithm works in rounds; during each round, each unique sequence of stops visited by a vehicle is explored at most once. RAPTOR is an efficient way to compute paths from an origin to all destinations. Additionally, the range-RAPTOR extension described in the original paper allows us to efficiently perform queries over the entire departure time window. This extension works by first performing a search at the last minute of the time window. We can then use the travel times to each transit stop in the network from that search as upper bounds on a search departing at the previous minute, because the earliest arrival departing at a minute cannot be later than the earliest arrival departing at the next minute. This method finds not a single path, but the set of paths that are optimal at some point in the time window. This means that our accessibility calculations assume that people have perfect information and choose the optimal path to reach their destination given their departure time.

Because the quality of the street network is a major determinant of whether people can even reach transit *(14, pp. 59ff)* our model includes a street network based on data from OpenStreetMap. We perform our pathfinding in several steps. First, we search on the street network using a standard Dijkstra algorithm, finding transit stops within a reasonable walking distance of the origin, as well as any direct paths to the destination that do not involve transit. We also precompute all possible transfers between stops using the same network and algorithm. We then perform the transit search using our variant of RAPTOR.

After every minute, we perform street network searches from the stops reached in the RAPTOR algorithm all the way to our grid of destinations. It is necessary to perform a search all the way to the destinations at every minute, rather than aggregating the times at transit stops and performing the street search at the conclusion of the transit search, because depending on the departure minute it may make sense to use a different transit stop to access the destination. For example, if you leave at 7:10 AM it may be best to take the train, but if you leave at 7:45 AM there may be an express bus that is faster and arrives at a different stop; one would lose this detail if travel times to each transit stop were aggregated together over the time window.

WORKING WITH UNDERSPECIFIED SCENARIOS

For the reasons outlined above, it makes sense to use a schedule-based network whenever possible. However, one often only knows the frequencies of lines that are added or modified in a scenario. We don’t want to switch to a purely frequency-based approach, because we are often modeling small changes to an existing network and wish to preserve the schedules of the routes that remain unchanged. Additionally, we wish to maintain the aforementioned attractive properties of the schedule search.

We handle frequencies by taking a Monte Carlo approach wherein we generate a large number of random schedules for the added or modified lines. It is not computationally feasible to perform a full range-RAPTOR search for each of these schedules, so instead we randomize the schedules at each departure minute of a single range-RAPTOR search. This means that the range-RAPTOR approach of bounding each search with the results of a search at a later minute is no longer applicable. In order to take advantage of the computational benefits of range-RAPTOR while retaining correctness, we first run a classic range-RAPTOR search at each minute using only the scheduled network (the portion of the network that is not assigned new random schedules at each minute). We then take the output of that search and use it to bound a search on the full network including the randomized schedules. The travel time using a subset of the routes in a network is an upper bound on the travel time of using all of the routes. It is important to include all lines, not just the frequency lines, in the latter search, to allow accessing a scheduled line using a frequency line. When the number of frequency lines is small, this approach results in minimal extra computation over the scheduled search.

We generally perform approximately 1000 Monte Carlo draws at each origin, spread evenly among the minutes of the departure time window; in the case study below we measure the variation in accessibility between separate runs of the algorithm. We currently assume that vehicles on a route arrive with exactly the specified headway, and the randomization determines their phase with respect to other routes. However, we could easily switch to an alternate method of producing random schedules.

Our initial instinct was to take a draw from the headway distribution independently each time the algorithm boarded a transit vehicle, rather than generating complete randomized schedules. However, the time a vehicle passes one stop is strongly correlated with the time it passes the previous stop on the same line. If we ignore this our results will be strongly positively biased, because the optimization algorithm will find the best boarding time at any nearby stop on the same route, effectively using the minimum of several random draws as the wait time.

There is much precedent for computing accessibility by simply assuming the boarding time of a route is half its headway *(e.g. 15)*. The advantage of our approach is that it produces a distribution of accessibility results for a range of possible schedules meeting scenario constraints, rather than simply computing an expectation. Additionally, our approach addresses the common lines problem mentioned above; when several (combinations of) transit lines are competing to get a customer to the same destination, half-headway is not an accurate representation even of the expectation. If several transit lines connect the origin and destination with, say, 30-minute headways, the expected wait is less than 15 minutes unless all the vehicles are coordinated to arrive at the same time.

Randomizing schedules reflects scheduling practice at many transit agencies. Schedules are often based on factors not correlated with passenger experience, such driver shifts or ensuring sufficient time and restrooms at layovers. However, sometimes lines are intentionally synchronized, e.g. buses that are timed to leave just after trains arrive, or a pulse *(14, pp. 164ff)*. We represent this by creating a dependency graph before performing our Monte Carlo process, and specifying that certain lines leave a stop exactly t seconds after another vehicle. This allows randomizing the unknown parts of the schedule while maintaining schedule synchronization.

We can also compute hypothetical minimum and maximum bounds on the accessibility by running a search at each departure minute with 0 or full headway boarding time on all frequency lines. However, these bounds are not tight because there is no way to schedule a transit system such that all users experience 0 or maximum wait times at all times.

In the case where a network is being created from scratch and all lines are defined using frequencies, we can defensibly ignore the departure time window, and instead run a search departing at a single minute on a large selection of random timetables, allowing the same approach to be used without being computationally prohibitive.

WORKING AT HIGH SPATIAL RESOLUTION

Unlike many transportation analysis systems, we represent origins and destinations on fine regular grids rather than using administrative boundaries or arbitrary polygons.

Traffic models traditionally aggregate trip endpoints into traffic analysis zones. However, spatial aggregation of data into arbitrary zones can have a strong, unpredictable effect on result quality. These scale and aggregation effects are known as the Modifiable Areal Unit Problem (MAUP). Openshaw (*16, p. 37)* concluded that all methods of analysis using spatially aggregated data would likely be affected by the MAUP, and that results of spatial analyses inherently depend on the zone system in use. The problem is not that results are necessarily incorrect or insignificant, but that it is impossible to know how much error is present. Because the problem is essentially unavoidable, it has become commonplace to ignore its existence and hope that ad hoc zoning systems will produce meaningful or interpretable results *(16, p. 31).*

The use of large zones in transportation planning is mostly an artifact of older technology with limited memory, storage, and computational power. An increasing number of datasets are available where each record represents a single individual, building, or city block. While some important data are only available at the neighborhood or municipal level (including output from many travel or land use models) it is not desirable to aggregate more detailed data sets to fit the largest common zone system. Dasymetric techniques exist to rationally allocate spatially aggregated data to locations where jobs, people, or facilities are most likely to be located, avoiding empty or unpopulated areas. Cadaster data sets containing the size or floor area of each individual building are often useful here.

Given advances in technology, we have adopted what Openshaw refers to as the convenient non-geographic approach: selecting zoning systems independently of the phenomena to be studied. We could represent each object individually, but this results in an unnecessarily large or even computationally prohibitive number of entities. Instead we aggregate to a fine raster grid. We situate opportunities with roughly as much spatial resolution as we expect in the temporal dimension. We report travel times and accessibility indicator thresholds in minutes; because walking is the slowest mode, the grid spacing should reflect the distance one can walk in about one minute. At our standard speed of 1.3 m/sec, this gives a spacing of 78 meters. While we often use somewhat larger cells, this distance serves as our guideline. Many countries now provide census data on similar grids *(17)*.

It is also important to ensure a lack of spatial correlation between the grid and the objects of study. A grid aligned to the compass points can introduce systematic error in places with similarly aligned rectangular street grids, especially when one grid size is a multiple of the other. ¼ mile street grids are common in the US midwest, and ¼ mile is approximately 400 meters. To avoid this problem, we use plane coordinate systems whose cell sizes are rarely round numbers.

Wong argues that researchers should perform sensitivity analyses of geographic resolution *(18)*, and due to our use of regular grids we can also apply concepts from sampling theory. Our future work should include analytic estimation of error margins and sensitivity analysis in which we increase the resolution of the grid until further increases can be shown to have a negligible effect on results.

A STATISTICAL VIEW OF SCENARIO EVALUATION

Figure 1 reveals the level of detail afforded by our method. These plots show cumulative accessibility to jobs as a function of travel time for a single origin point. Job accessibility is on the vertical axis (square root scale) and travel time increases to the right (linear scale). Similar plots can be made for every origin.

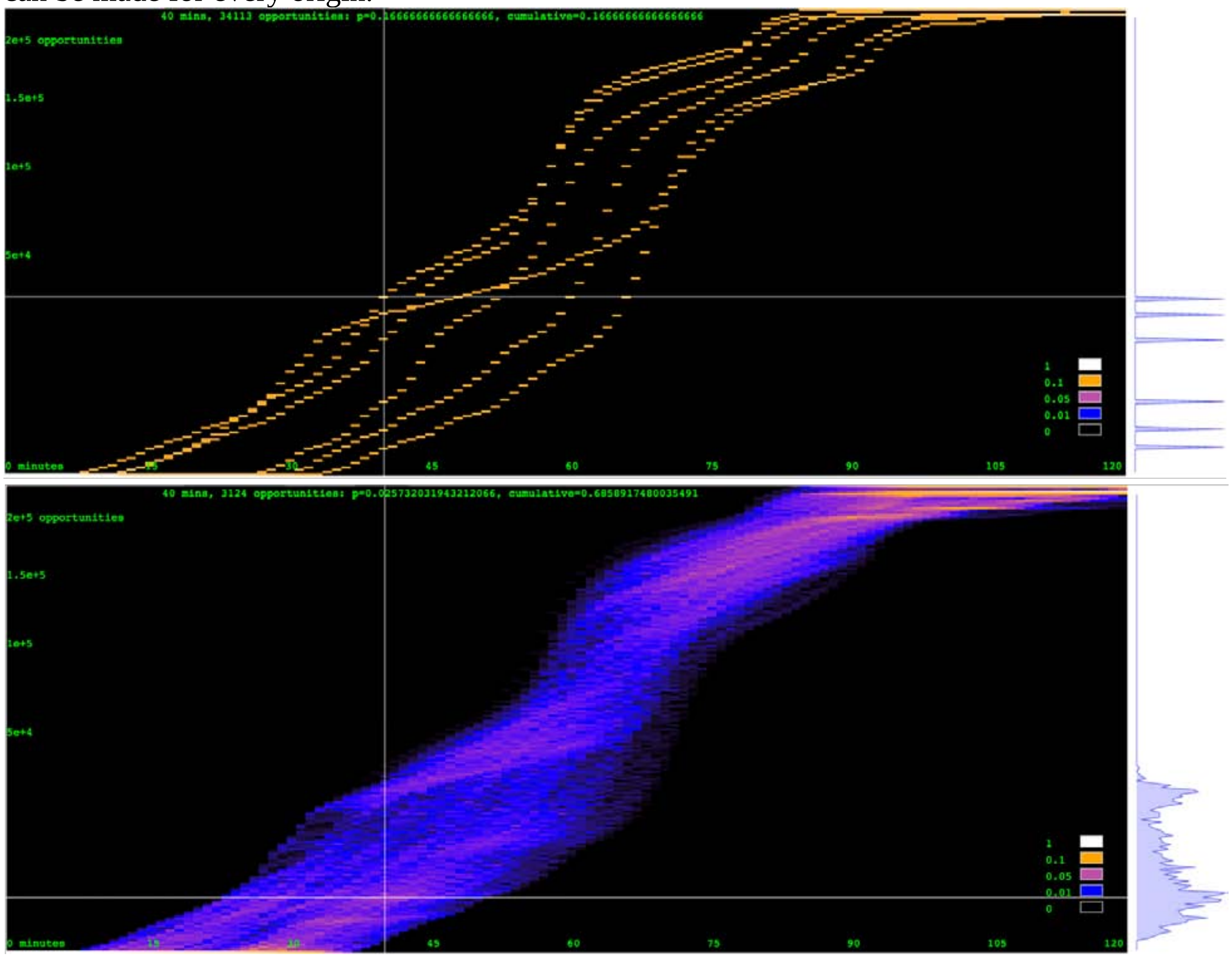

Figure 1: Comparing accessibility distributions for 6 and 1000 randomized schedules

Each departure time within the analysis window and each randomized schedule yields different travel times to destinations throughout the network and therefore different accessibility results. Thus, for a given origin point and travel time the cumulative opportunities indicator does not have a unique value. The upper plot shows only six cases, and the six distinct cumulative accessibility curves are clearly visible. With the 1000 random schedules in the lower plot, the curves overlap and we obtain a continuously varying density.

The margins of these plots show vertical slices at a travel time of 40 minutes. In the upper case, six distinct values ranging from 700 to 34,000 are visible with equal weight. In the lower case we obtain something resembling a probability distribution. Under this scenario, depending on when exactly the passenger leaves home and whether the routes that passenger relies upon are in or out of phase at transfer points, the 40-minute job accessibility is bimodal with strong peaks around 3,400 and 37,300 reachable jobs (an order of magnitude apart). For a given travel time we can of course compute mean or median accessibility, but such a summary measure is far from telling the whole story. In fact, in this case a passenger is unlikely to ever experience that accessibility value.

This highlights a challenge in comparing scenarios. Unless the accessibility distributions for two scenarios are disjoint, one cannot be certain that one scenario would yield superior accessibility to the other in practice. This is not a shortcoming of our methods, it is an inherent part of any effort to measure the accessibility characteristics of an incompletely specified (headway-based) network or consider rider experience over a range of departure times. However, since we treat our results as distributions, we can use a hypothesis test to compute a probability that one scenario is an improvement over another.

APPLICATION TO THE METROPOLITAN AMSTERDAM REGION

In late 2015, the authors supported Dutch Infrastructure and Environment Ministry in the *OV Toekomstbeeld 2040* (Public Transit Future Vision 2040) which aimed to establish a methodology for envisioning spatial development, including both transportation and land use components *(19)*. This pilot project focused on the “south wing” (*zuidvleugel*) of the heavily urbanized, polycentric Randstad region of the Netherlands, centered on metropolitan Rotterdam and The Hague. Conveyal and Movares collaborated with Dutch firms APPM and Goudappel Coffeng as well as Dutch Railways (NS), regional transit operators, and municipalities to elaborate a range of four transportation scenarios, ranging from severe transit cuts to ambitious expansion and service reinforcement. The techniques described in this paper were chosen as a way of assessing the accessibility impact of collaboratively designed scenarios. This was one of the first major applications of these accessibility methods in an official planning process, allowing us to refine them based on feedback and experience.

The methodology developed during this pilot project was approved by the Ministry, and all other Dutch urban regions were asked to replicate and build upon it. The authors then collaborated on a second iteration of this process for the northern Randstad, centered on the Metropolitan Amsterdam Region, referred to as MRA in Dutch. Here the process was revised based on experience in the South Randstad.

The authors favor running regional analyses for every point in a regular grid, but for consistency with the pilot project a different approach was adopted. We used an automatic process to choose candidate points that were local maxima of population and employment density derived from job and population counts at the individual building level from the LISA database *(20)*. These were then combined with notable locations such as universities. The list was culled in stakeholder meetings to 188 analysis points, referred to as the *magneten* (magnets).

The process of magnet selection was then repeated using future data derived from the Amsterdam VENOM traffic model *(21)* and the Netherlands Regional Model *(22)*. Employment and population growth factors derived from these models’ neighborhood-sized zones were applied to building-level Census data and new magnets were found; the same building-level projections were also aggregated to produce high-resolution destination density grids for these various years. The current and future magnets were compared and a single set was chosen for both baseline and future analysis. The number of inhabitants and jobs within 45 minutes was then calculated for each magnet in each study year using our accessibility techniques. In the final reports, these measures were referred to as *netwerkkwaliteit* (network quality), which is somewhat of a misnomer since accessibility measures the location-specific effectiveness of the combined transportation and land use system.

Unlike the South Randstad pilot project, in metropolitan Amsterdam every modification to the transit system was tested separately before being integrated into complete scenarios. 33 micro-scenarios were created, each one representing an individual measure: a change in frequency or speed, a change to a route alignment, or the elimination or creation of a route. Variations were explored such as alternative alignments or transfer points. Accessibility indicators were calculated at all the magnets under each micro-scenario for comparison against the baseline network and against one another. This allowed planners to weigh the accessibility impact of each separate measure, including network effects, iteratively assembling them into complete packages that would become the final scenarios.

Eventually a final scenario emerged from discussions. This scenario was generalized into broader statements about which investments MRA would like to make, accompanied by cost estimates. Along with similar results from the South Randstad and other parts of the Netherlands, these regional conclusions will be taken into account in the formulation of a larger national plan that determines which projects will be realized and how funding will be allocated.

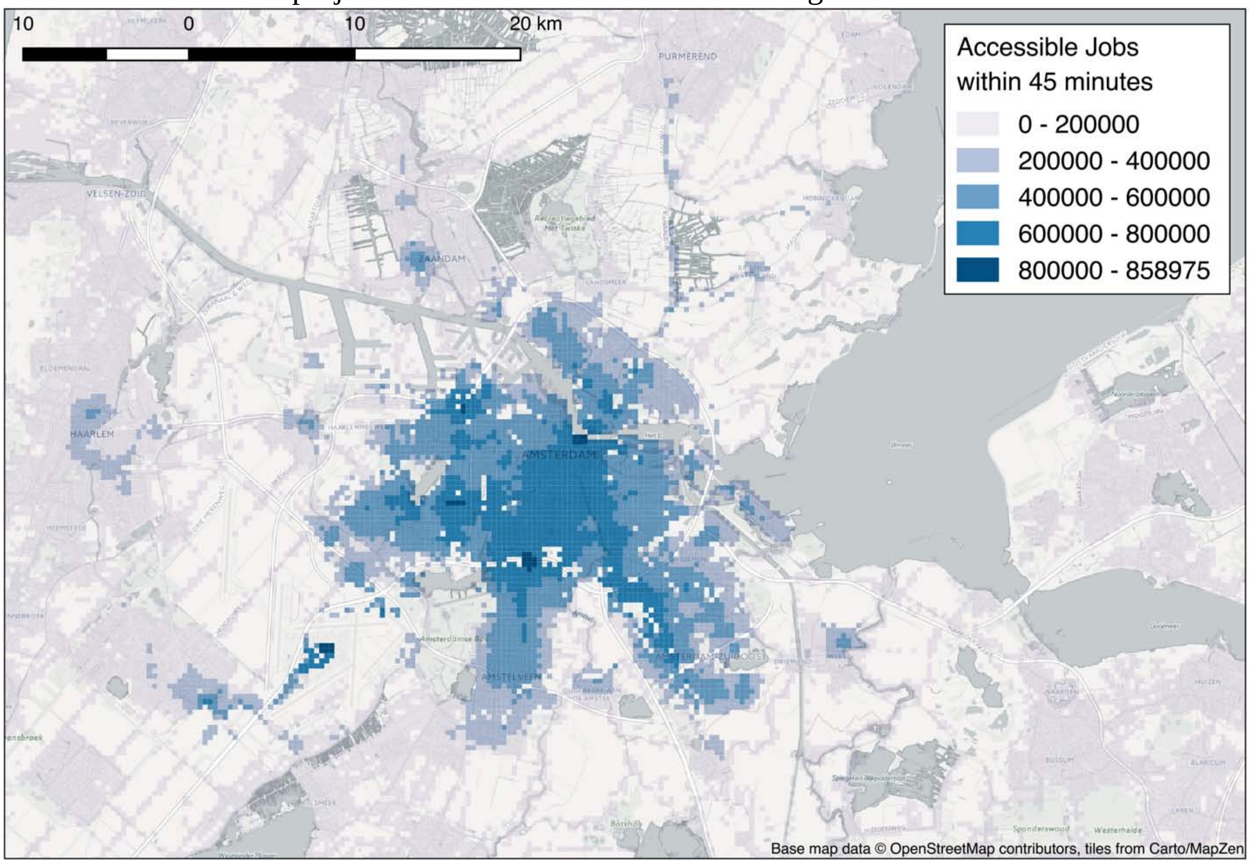

Figure 2: Baseline job accessibility, MRA

Regional Results and Interpretation

The figures show the results of a complete accessibility analysis using a roughly 200-meter raster. Figure 2 shows the number of jobs reachable within 45 minutes of each grid cell using the current network. For this project, we used the mean rather than the median travel time to compute accessibility. This should introduce minimal bias as the network is well-connected; most trips are possible throughout the time window. Figure 3 shows changes in this indicator across the region under the final scenario. In both images, land use is set to 2030 levels to highlight the effect of the transportation network alone.

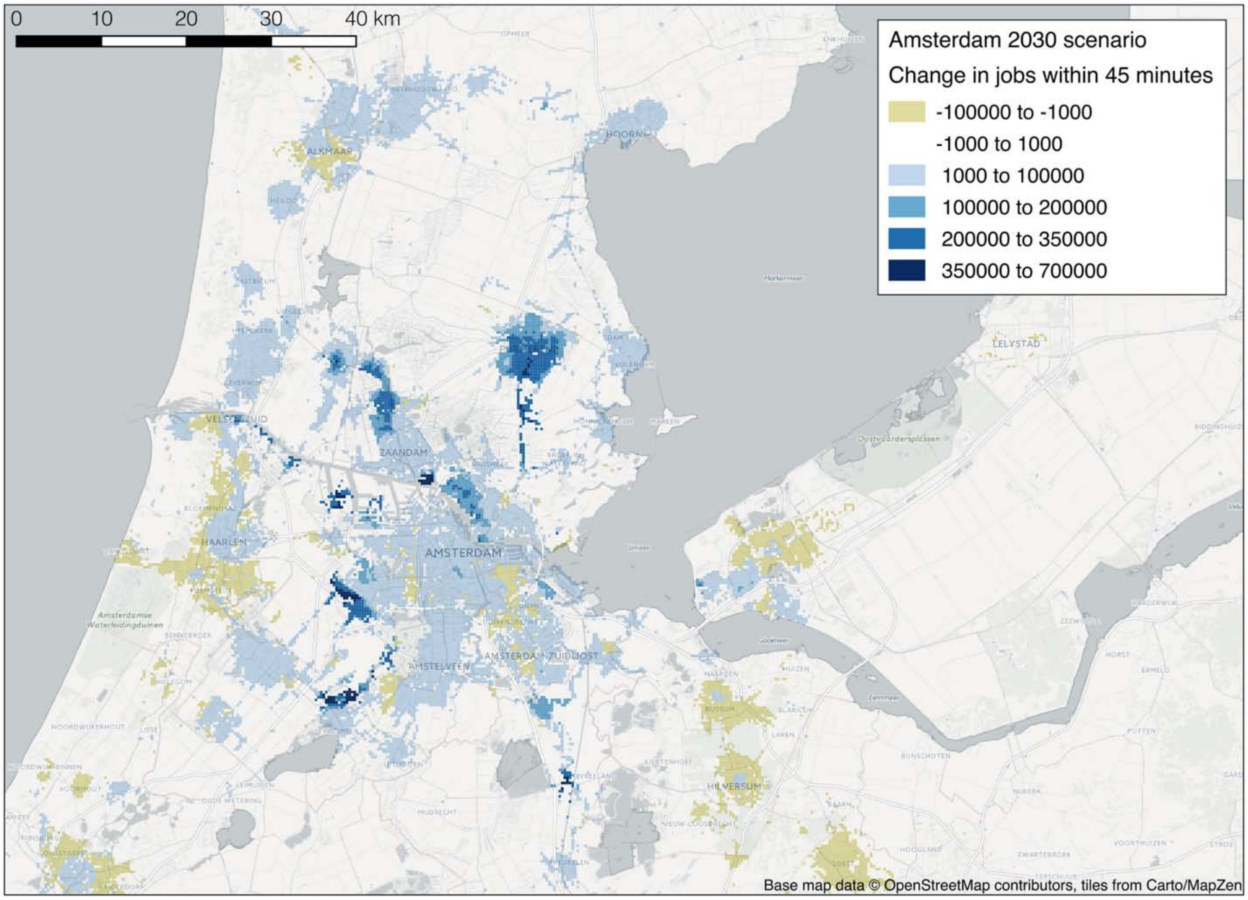

Figure 3: Change in job accessibility, MRA

Exploring results of this kind for both the combined scenario and its component micro-scenarios allowed planners to come to a number of conclusions about future transport policy in the Amsterdam region. For example, planners found that removing all bus lines that run less than four times an hour had surprisingly little effect on accessibility throughout the region. Though some small towns were cut off from public transport, the system as a whole remains coherent, and the effect on accessibility at the "magnet" analysis locations was slight. There was significant impact in the more rural regions, but even there only a 10-15% decrease was observed because some frequent buses remained. Infrequent direct service that was removed was mostly compensated for by frequent services requiring transfers *(see also 14, pp. 150ff).* When and if another, cheaper option emerges to keep people in smaller towns and villages mobile, these results could support replacing expensive infrequent lines with that cheaper mode.

The analysis results strongly supported the introduction of a tram-train mode, combined with the segregation of traffic into light and heavy rail in the congested tunnel between Amsterdam and Schiphol airport. This tunnel has two tracks per direction, all of which are currently used by heavy rail and near capacity. It would be extremely expensive to enlarge, as it runs near underground parking garages and airport facilities. Scenarios were tested in which all *Sprinter* local service was replaced by extending Amsterdam metro and tram lines to serve local rail stations. This strategy, termed *regiorail* (regional rail), would allow allocating two tracks in the tunnel to light rail and the other two to intercity heavy rail, causing a threefold increase in capacity on the light rail tracks due to closer spacing of vehicles and shorter dwell times at stops. This plan had already been proposed as a way to increase capacity, but our analysis also showed a significant increase in accessibility because transfer waiting time was eliminated for trips to and from villages and neighborhoods away from major rail stations.

One scenario demonstrated the value of considering accessibility results within a public debate process rather than following them blindly. Almere is a largely residential town east of Amsterdam. Given the high price of housing in central Amsterdam, it is widely seen as less expensive “overflow” housing for workers who still need to access the city center. Housing capacity there is set to grow over the study period as reclaimed land is opened to construction. A scenario tested the effects of extending an existing rail line from Amsterdam through Diemen directly across the IJmeer lake toward Almere. It cut travel times significantly and showed very strong improvements in accessibility in key places, but the realization of such a project is not feasible: this lake is a cherished and protected natural area, and the very long bridge that this scenario implies is unacceptable because of environmental damage. The alternative of building a tunnel would also be environmentally questionable and even more expensive. Both options attach excessive costs or externalities to accessibility benefits.

Result Stability

As described above, scenarios using headways are analyzed by generating many random schedules. Accessibility values derived from average travel times are then extracted to account for diverse travel times that might be encountered once the schedule is realized and experienced at different departure times. It is not feasible to exhaustively test the vast number of different schedules that fit a headway-based scenario description; randomization allows us to quickly sample enough schedules to build a statistical picture of the scenario's probable impact.

This use of randomness does mean that travel times and therefore accessibility values will differ slightly from one run to the next. The more random schedules we generate, the better accessibility results will approximate the distributions that would be found by exhaustively exploring every schedule. For a given scenario, it is important to understand how stable results are, how far they may be from the true values, and how this source of error diminishes when we generate more schedules.

To demonstrate this issue, we re-ran a complete regional analysis of the final metropolitan Amsterdam scenario four separate times, using a rather low number of 240 randomized schedules each time; we use approximately 1000 for final analyses. We then calculated the coefficient of variation for mean 45-minute job accessibility at each origin point in the region over the different runs. Figure 4 shows the magnitude and spatial distribution of this randomization error.

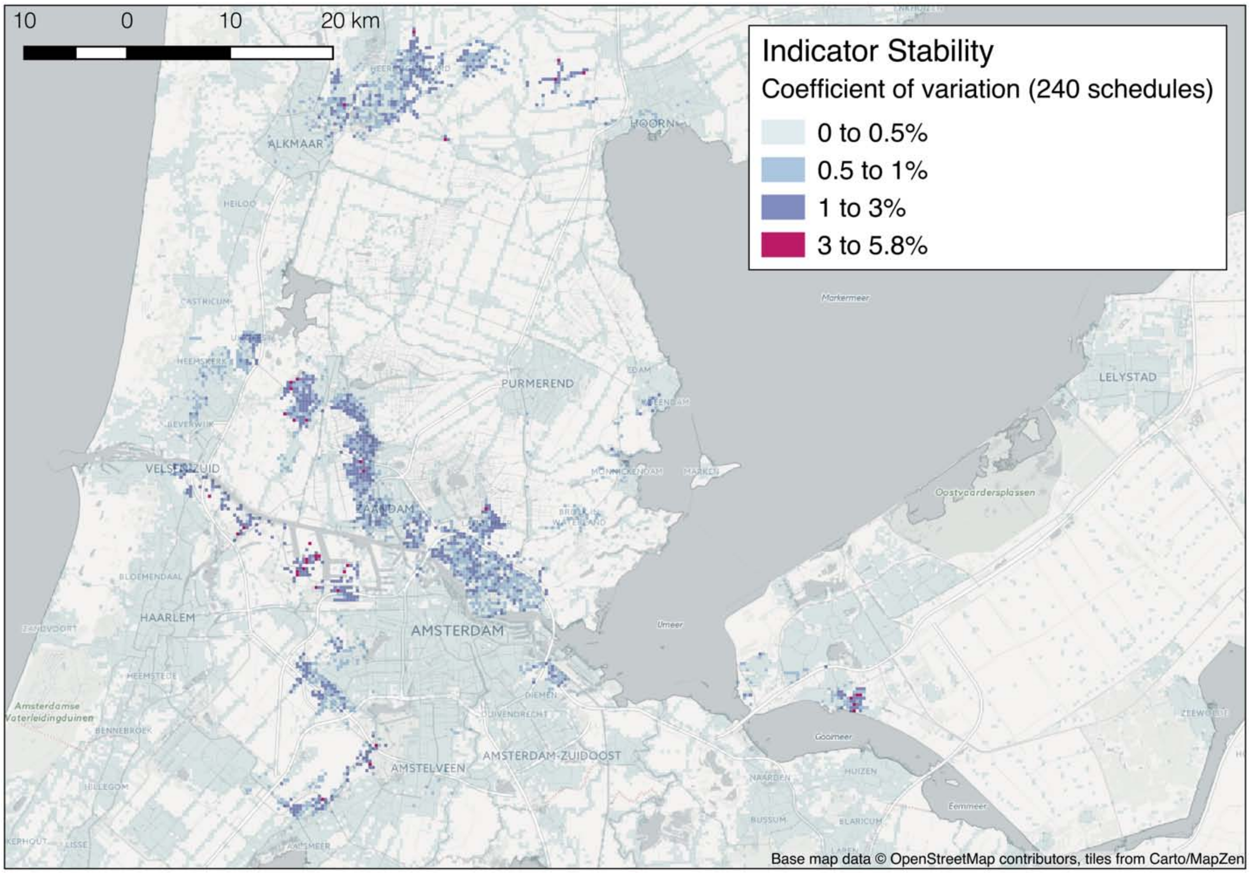

Figure 4: Coefficient of variation of mean accessibility

When evaluating this scenario with 240 schedules, accessibility results are generally quite stable with a CV below 0.5% in most locations. However, the error shows distinct spatial patterns with CV's reaching 3% to 6% in isolated locations. If more uniformly accurate results are desired, the number of randomized schedules would need to be increased and the analysis re-run until the dispersion falls below the desired level at all points.

CONCLUSION

Because of our system's use of open standard data formats, we were able to quickly build a baseline network model from open GTFS and OSM data that was more current and precise than existing transport models in the region, even reflecting existing transfer timings and barriers to pedestrian or bicycle movement such as waterways. Thanks to our stochastic schedule-based approach, we were able to maintain that level of detail on the part of the network that was untouched while patching in speculative changes and understand the level of uncertainty in results due to a headway-based description of those changes.

Planners were able to visualize and discuss the impact of each potential change to the network individually before assembling them into final scenarios. The accessibility impact of these measures was visible to them throughout the process. Our optimized and parallelized implementation means that it is now technically feasible to provide continuous, immediate, and nuanced feedback on regional accessibility impact in near real time while scenarios are still being sketched out.

REFERENCES

1. Brömmelstroet, M. te, C. Silva, L. Bertolini. *COST Action TU1002 – Assessing Usability of Accessibility Instruments*. COST Office, April 2014. http://www.accessibilityplanning.eu/wp-content/uploads/2014/05/COST-REPORT-II.pdf.
2. Geurs, K. T., and B. van Wee. Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography*, Vol. 12, No. 2, 2004, pp. 127–140.
3. El-Geneidy, A. M., and D. M. Levinson. Access to Destinations: Development of Accessibility Measures. Report MN/RC-2006-16. Minnesota Department of Transportation, 2006. http://conservancy.umn.edu/handle/11299/638.
4. Owen, A., and D. M. Levinson. Access Across America: Transit 2014 Methodology. Publication CTS 14-12. University of Minnesota, Minneapolis, 2014. http://www.its.umn.edu/Publications/ResearchReports/pdfdownload.pl?id=2504.
5. Busby, J. R. Accessibility-Based Transit Planning. Massachusetts Institute of Technology, Cambridge, MA, 2004. http://hdl.handle.net/1721.1/32414
6. Geurs, K. T. Accessibility, Land Use and Transport: Accessibility Evaluation of Land-use and Transport Developments and Policy Strategies. Utrecht University, Utrecht, 2006. http://igitur-archive.library.uu.nl/dissertations/2006-0613-200112/
7. Palmateer, C., A. Owen, and D. M. Levinson. Accessibility Evaluation of the Metro Transit A-Line. University of Minnesota, Minneapolis, 2016. http://conservancy.umn.edu/handle/11299/180900.
8. Liu, Y., J. Bunker, and L. Ferreira. Transit Users' Route-Choice Modelling in Transit Assignment: A Review. *Transport Reviews*, Vol. 30, No. 6, 2010, pp. 753–769.
9. Chriqui, C., and P. Robillard. Common Bus Lines. *Transportation Science*, Vol. 9, No. 2, 1975, pp. 115–121.
10. Spiess, H., and M. Florian. Optimal Strategies: A New Assignment Model for Transit Networks. *Transportation Research Part B: Methodological*, Vol. 23, No. 2, 1989, pp. 83–102. https://www.researchgate.net/publication/222385476. Accessed July 23, 2016.
11. Nguyen, S., and S. Pallottino. Equilibrium Traffic Assignment for Large Scale Transit Networks. *European Journal of Operational Research*, Vol. 37, No. 2, 1988, pp. 176–186. https://www.academia.edu/11918753/Equilibrium_traffic_assignment_for_large_scale_transit_networks_1. Accessed July 27, 2016.
12. Citilabs. *Multi-modal accessibility methodology & application*. 2016. http://cdn2.hubspot.net/hubfs/509866/SUGAR-ACCESS-White-Paper-v1-5_01.pdf?t=1453989134670. Accessed Aug 1, 2016.
13. Delling, D., T. Pajor, and R. Werneck. Round-Based Public Transit Routing. 2012. http://research.microsoft.com/pubs/156567/raptor_alenex.pdf.
14. Walker, J. *Human Transit*. Island Press, Washington, DC, 2012.
15. Lee, B. H.-Y. Parcel-level Measure of Public Transit Accessibility. In *Access to Destinations* (D. M. Levinson and K. J. Krizek, eds.), Elsevier, Amsterdam, pp. 215–239.

16. Openshaw, S. The Modifiable Areal Unit Problem (CATMOG 38). Geo Books, Norwich, 1984.
17. European Forum for Geostatistics. *GEOSTAT 1A – Representing Census data in a European population grid*. Version 8.0, 2012. http://ec.europa.eu/eurostat/documents/4311134/4350174/ESSnet-project-GEOSTAT1A-final-report_0.pdf.
18. Wong, D. The Modifiable Areal Unit Problem. In *SAGE Handbook of Spatial Analysis* (A. S. Fotheringham and P. A. Rogerson, eds.), Sage, Los Angeles, pp. 105–123.
19. APPM Management Consultants and Goudappel Coffeng. Toekomstbeeld OV: Pilot Zuidelijke Randstad. 2016. http://mrdh.nl/system/files/vergaderstukken/I.02.3_Bestuurlijke%20Rapportage%20Toekomstbeeld%20OV%20Zuidelijke%20Randstad_1.pdf. Accessed August 2, 2016.
20. LISA Foundation. LISA-Handboek: Definities, protocollen en achtergronden van LISA, March 2015. https://www.lisa.nl/include/LISA_Handboek_versie_maart_2015.pdf.
21. Stadsregio Amsterdam. Regional Verkeersmodel VENOM. https://www.stadsregioamsterdam.nl/pagina/20160104-regionaal-verkeersmodel-venom. Accessed July 21, 2016.
22. Joksimovic, D., and R. van Grol. New Generation of Dutch National and Regional Models—An Overview of Theory and Practice. *European Transport Conference*, 2012.