

SCHOOL OF SUSTAINABLE ENGINEERING AND THE BUILT ENVIRONMENT



**Personal Environmental Exposure Digital Twin  
Icarus 25 Framework**

Sneha Roy and Mikhail Chester

ASU-METIS-25-TRS-003

November 2025

TECHNICAL REPORT

# PERSONAL ENVIRONMENTAL EXPOSURE DIGITAL TWIN ICARUS 25 FRAMEWORK

**Sneha Roy and Mikhail Chester**

Metis Center for Infrastructure and Sustainable Engineering  
School of Sustainable Engineering and the Built Environment  
Arizona State University

November 2025

Report No. ASU-METIS-25-TRS-003

## **ABSTRACT**

Icarus is an open-source urban digital twin that links a full metropolitan activity-based travel model with high-resolution environmental data to measure and minimize personal heat exposure. By embedding sub-meter, hourly Mean Radiant Temperature ( $T_{MRT}$ ) rasters into a multimodal street network, the framework both captures how individuals accumulate exposure during their daily travel, as well as is capable of rerouting them to cooler alternative paths. Earlier applications showed that  $T_{MRT}$ -weighted routing can cool up to ten times more trips than infrastructure retrofits alone, positioning Icarus as a decision-support tool for planners, public-health officials and researchers concerned with climate equity.

Significant improvements to the earlier version of the Icarus framework are documented in this report. In addition to incorporating multiple hazards with the objective of minimizing each, this new framework also accounts for changing hazard conditions as people complete their trips, combining improved data resolution with more adaptive routing methods. Key upgrades also include a time-dependent routing engine that accounts for shifting hazard patterns during each trip and finer spatial precision through edge-splitting at trip endpoints. The framework adds sub-edge temperature sampling and localized network indexing, greatly improving both accuracy and computational efficiency. Together, these developments turn Icarus 25 into a dynamic, multi-hazard digital twin that can measure and minimize personal exposure across millions of walking and biking trips while remaining deployable for large-scale urban analyses.

## TABLE OF CONTENTS

1. Background & Motivation .....	3
2. Urban Digital Twin .....	3
3. Core Workflow of Icarus .....	4
4. Data Sources.....	5
4.1. Environmental Hazards .....	5
4.2. Transportation Network.....	6
4.3. Land Use .....	6
4.4. Travel Behavior Model .....	6
5. Core algorithms .....	7
5.1. Edge-Level Hazard Attribution .....	7
5.2. Heat-Exposure Objective and Routing .....	8
5.3. Time-Dependent Dijkstra Routing .....	8
5.4. Outputs .....	8
6. Time-Dependent Dijkstra for Single and Multi-Hazard Exposure .....	9
6.1. Dynamic vs. Static (Snapshot) Routing .....	9
6.2. Multi-Hazard Extension - Minimizing Multiple Simultaneous Hazards .....	10
7. Additional routing enhancements .....	13
7.1. Spatial Precision at Trip Ends: Edge Splitting & Temporary Nodes .....	13
7.2. Guarding against backtracks to split points.....	14
7.3. Localized Subgraphs & Spatial Indexing .....	14
References .....	15

## 1. BACKGROUND & MOTIVATION

The Icarus framework was developed to address the challenge of estimating personal heat exposure. The framework was initially conceived in 2015 and with support from several funding agencies (namely the U.S. National Science Foundation) Icarus was developed. The first alpha release provided the foundation for assessing heat exposure of active mobility travelers using Phoenix, Arizona as the test bed. The framework was one of the first to show how *individuals* accumulate heat exposure (Li et al. 2023), how exposure can be reduced through behavioral and infrastructure change (Li et al. 2023), and how we can shift our understanding of negative thresholds by considering both outdoor heat exposure and indoor air conditioning reprieve (Li et al. 2024).

For a system designed to understand and mitigate urban heat risks, Phoenix serves as an ideal testbed. The region spends nearly half the year above 32 °C (90 °F), with NOAA (2022) noting more than 170 such days annually. In 2021, 339 deaths in Maricopa County were attributed to heat exposure (MCDPH, 2022). Climate research further warns that future heat waves will grow more intense and persist for longer (IPCC, 2021; Ebi et al., 2021). At the same time Phoenix has one of the most advanced co-governance efforts to reduce negative health outcomes (MCDPH, 2022).

Most heat-exposure studies look at groups of people in fixed locations, often assuming their exposure reflects the conditions around their homes (Guzman-Echavarria et al. 2022) or workplaces. Icarus, on the other hand, accumulates exposure estimates by simulating individual active pedestrian and bike trips over the course of a day. It accounts for the exposure accumulated by agents in an activity-based model (ABM) over all the trips they make during the day and the time spent at hazard-exposed destinations before embarking on subsequent trips. Finely resolved hazard rasters help answer questions like how much heat exposure do active travelers really accumulate across the region? To what extent are current exposure estimates accurate or biased by the resolution of the inputs? How can such insights be mobilized to inform cooling interventions and resilient urban design (Gunawardena et al. 2017; Li et al., 2023)?

Here, we describe the advancements to the initial Icarus model in (Li et al. 2023). We refer to these advancements as Icarus 25. Icarus 25 has been developed with advanced spatial analytical techniques, new data inputs, and the ability to assess concurrent hazards (e.g., heat and air quality). The rest of this report focuses on both describing the core logic of Icarus 25 and the new techniques from past versions of the framework.

## 2. URBAN DIGITAL TWIN

Icarus is an urban digital twin that links a synthetic population's daily travel patterns with high-resolution geospatial data layers to quantify, and ultimately mitigate, the environmental exposures experienced by active mobility travelers. Icarus integrates diverse spatial datasets to characterize land use across the study area, while drawing on output from an activity-based travel demand model (ABM) to determine where travelers go and why. Although the ABM produces trips at the level of Metropolitan Analysis Zones (MAZs), Icarus refines these outputs by generating parcel-level origin and destination points while preserving person-level trip chaining. This step keeps travel

behavior realistic. For instance, work trips terminate at parcels aligned with occupational land use, school trips at the nearest school parcel within the assigned MAZ, and shopping trips at retail parcels. Earlier applications of the framework (Li et al. 2023) used this approach to estimate Mean Radiant Temperature (MRT) exposure for about 625,000 active trips in Phoenix and to evaluate the cooling potential of different behavioral and infrastructure strategies. Companion studies (Li et al. 2023; Li et al. 2024) showed that the framework can adapt from detailed, trip-level analysis to broader assessments of heat exposure across traveler groups or the region as a whole. By modeling how pedestrian–infrastructure interactions evolve over the day in response to environmental stimuli such as heat or particulate matter with almost streetside precision, Icarus effectively operates as an urban digital twin that complements traditional climate and exposure modeling to provide potentially actionable insights.

### 3. CORE WORKFLOW OF ICARUS

The core workflow of Icarus 25 is shown in Table 1.

*Table 1. Major modules comprising the workflow of Icarus 25.*

<b>Module</b>	<b>Function</b>	<b>Key Inputs</b>
<b>Agent Definition</b>	Import the synthetic population from a metropolitan-scale ABM. Each agent arrives with a tour schedule, activity type, start time, and origin/destination at the Maricopa traffic-analysis-zone (MAZ) level.	Activity Based Model Input and Outputs: Land Use, Trip types, characteristics of agents and tours.
<b>Parcel-level granularization:</b> Redistribution of trip Origins and Destinations	Crosswalk every household and activity type in the ABM to parcel-level land-use (e.g., home, school, retail, recreation), based on person and trip activity type, for example, school trips from students get assigned to school-type parcels, work trips get assigned to a pre-determined parcel that is related to occupation type of person, and so on. We thus impute a deterministic parcel for each trip start and end point, replacing the MAZ centroids with a geocoded address or building footprint.	County or municipal parcel GIS datasets with land-use attributes
<b>Network Synthesis</b>	A multimodal pedestrian network derived from open-street geodata (Open Street Maps - OSM) is conflated to a hazard raster like heat or air quality indicator (PM <sub>10</sub> or PM <sub>2.5</sub> ). Each link in the network is assigned a hazard-cost by multiplying hazard-unit and link-length that changes hourly.	OSM extracts of local street- and pedestrian centreline files, hazard rasters
<b>Routing Engine</b>	Run customizable routing algorithms to find a path for each active passenger trip from its origin to destination: <ul style="list-style-type: none"> <li>• Minimum-distance (baseline).</li> </ul>	Pedestrian/Bike Routing Network

	<ul style="list-style-type: none"> <li>• Minimum-exposure using user-specified hazard rasters</li> </ul> Algorithms can be swapped between Dijkstra, A*, k-shortest and time-dependent exposure costs.	
<b>Exposure Estimation</b>	Compute cumulative exposure per trip and summary statistics per traveler, neighborhood, corridor or scenario.	Baseline (Shortest) Paths and exposure-minimized paths for trips
<b>(Optional) Scenario control</b>	Some example scenario tests: 1. Replace hazard raster with desired spatial weights or datasets at different resolutions, 2. Test or locate cooling corridors or new shade structures 3. Isolate and analyze different activity schedules.	Alternative hazard or network layers depicting scenarios to be analyzed

Icarus is designed to measure individual-level and link-level exposure as agents move across an underlying network. This underlying network can be resolved to any metric of exposure one wants to measure, for instance, heat, particulate matter, other air pollutants, or any hazard to people that can resolutely be expressed in time and space. In this study, we route active travelers, namely pedestrians, on a walking network. For each synthetic pedestrian that Icarus routes between the origin and destination, the model accounts for, and incorporates, the trip's activity type and the time of day during which it occurs. Spatial data are joined with this trip information including environmental temperature, pedestrian and biking infrastructure, and indoor/outdoor temperature of the area within the trip. This is done primarily for the following uses:

1. Measuring trip-level or link-level exposure for trips following typically projected routes.
2. Rerouting trips between their origin and destination following paths that minimize exposure, as opposed to the more commonly used routing approach – minimizing travel time (or distance).

**4. DATA SOURCES**

This section provides a short description of the data sources that Icarus 25 uses as its input.

**4.1. Environmental Hazards**

Temperature serves as the primary environmental hazard whose spatial resolution is systematically varied and evaluated for its effect on exposure measurement and routing outcomes. Temperature values are used both to determine routing weights (for minimum-exposure routing) and to calculate the exposure associated with each path post hoc. In one application of this model, different temperature resolutions were tested to assess their influence on routing decisions. Regardless of whether a scenario employed heat-aware or distance-based routing, post-routing exposure is computed for every trip to enable direct comparison of exposure savings between routing

strategies. This allows assessment of the “cost” savings achieved by taking a longer but cooler route. Furthermore, in multi-hazard routing scenarios (discussed in Section 6.2), post hoc exposure is calculated for all considered hazards, regardless of the objective being minimized and the method utilized to do so.

Mean radiant temperature ( $T_{MRT}$ ) is the composite radiative load on a human body, aggregating shortwave and longwave fluxes from all directions; it is a primary driver of outdoor thermal exposure and comfort in urban settings.  $T_{MRT}$  is commonly modeled with SOLWEIG, which resolves 3-D radiation fields at fine spatial scales, and it has been extensively measured in the field using the six-directional “MaRTy” platform developed by (Middel, 2016; Middel and Krayenhoff, 2019), who have shown how shade geometry strongly modulates  $T_{MRT}$  at pedestrian height (Middel et al. 2021). We therefore use  $T_{MRT}$  as the hazard variable for both exposure accounting and routing.

#### 4.2. *Transportation Network*

An underlying spatial network acts as the graph over which trips are routed. Each edge stores attributes invoked by alternative objectives (time, distance, exposure, or a combination of these). We use a combination of OpenStreetMap walk and bike layers (~ 1- 5m geometry), the granularity of which is within the range of and compatible with our highest resolution exposure dataset (1m).

#### 4.3. *Land Use*

Icarus 25 uses highly resolved land use information at parcel scale that considers trip activity types to impute origins and destinations from ABM to more precise parcels using these steps:

1. **MAZ seed** – Each trip starts with an ABM-supplied origin-MAZ and destination-MAZ plus activity type and departure time.
2. **Parcel lookup** – For every (MAZ, activity type) pair, we query the parcel database for parcels whose land-use matches that activity (e.g., “primary school” for education trips).
3. **Probabilistic assignment** – When multiple candidate parcels exist, a probability weight proportional to the parcel floor area is applied and one parcel is drawn as such.
4. **Coordinate extraction** – The parcel centroid (or an entrance point) becomes the precise origin or destination input to the routing engine. This is constant for each trip and does not change across model scenarios.

Pedestrians are then routed between these origins and destinations using paths optimized for minimizing exposure or distance.

#### 4.4. *Travel Behavior Model*

We ingest the MAG ABM day plan for ~3.8 M synthetic persons, representative of the real population and their travel patterns. Each record provides a tour (sequence of activities), the mode for each inter-activity leg, a departure time, and MAZ-level origin/destination. We retain legs that are (i) primary walk or bike trips and (ii) walk/bike access/egress to transit optionally (configurable). Icarus 25 focuses on purely active trips during 07:00–20:00 within the MRT spatial footprint.

## Parcel localization and maintaining trip chains observed in ABM

ABM MAZ endpoints are granularized to parcels using the land-use crosswalk described earlier. This is done to preserve escorting and other chain dependencies by keeping tour ordering intact while replacing zone centroids with curb-front coordinates. Origins/destinations are fixed across scenarios such that only the hazard layer (MRT) and routing objective need to vary.

## 5. CORE ALGORITHMS

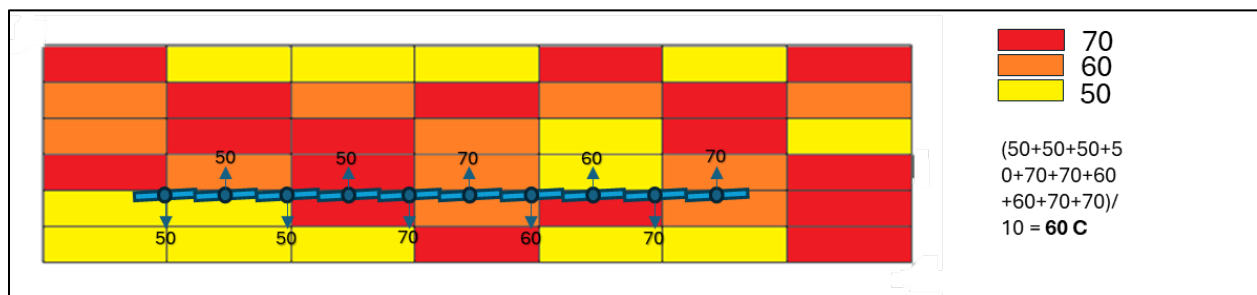
Icarus 25 converts environmental raster data into link-level exposure costs and then applies a time-aware routing procedure to identify least-exposure paths across the pedestrian network. Hazard values are first attached to edges, and then pedestrian trips are routed through this exposure-aware network.

### 5.1. Edge-Level Hazard Attribution

Urban exposure estimates, especially those measured at high resolutions (like MRT) have the potential to vary sharply even across a single street segment. Temperature or pollutant levels at one end of a block often do not match conditions at the other. Using a single raster value for an entire street segment can therefore misrepresent exposure. To reduce this bias without physically subdividing the network, and to simultaneously maximize using the resolution of the hazard rasters informing the exposure estimates, Icarus 25 samples each edge at ten points along its geometry:

- Parameterization - Each link's link geometry is mapped onto a normalized scale from 0 (start) to 1 (end).
- Transect Points - Ten equally spaced fractions (0.1 ... 1.0) define lookup positions along the link.
- Raster Sampling - At each point, the nearest raster cell (for  $T_{MRT}$  or any other hazard surface) is retrieved.
- Aggregation - The mean of these ten values becomes the representative hazard level  $H_e(h)$  for edge  $e$  at hour  $h$ .

This method captures sub-edge variation reasonably well without resolving the underlying network layer further.



## 5.2. Heat-Exposure Objective and Routing

The hazard cost for each edge is dynamic, varying by five-minute intervals and interpolated using hourly  $T_{MRT}$  inputs. During routing, the cost of each edge 'e' is calculated at the traveler's predicted entry time to that edge rather than calculating a snapshot of costs of all prospective links forming a path at trip departure (more in Section 6.1).

1. The hourly series  $H_e(h)$  is linearly interpolated into five-minute bins  $H_e(t)$  to capture ephemeral fluctuations in exposure estimates.
2. The edge cost combines hazard intensity and edge-traversal time:

$$C_e(t) = H_e(t) \times \Delta t_e$$

Where,

- $C_e(t)$  = total exposure cost of edge e when entered at time t;
  - $H_e(t)$  = interpolated hazard (e.g., temperature or pollutant concentration) on edge 'e' at time 't';
  - $\Delta t_e$  = travel time required to traverse edge e.
3. Optional safeguards, such as a small time penalty to encourage shorter travel trips as tie breakers for two equal cost paths and a limit on allowable detour length, are also implementable.

## 5.3. Time-Dependent Dijkstra Routing

Routing is carried out in a time-expanded space in which each state is defined by a pair of the network node and the time bin at which the traveler arrives there (node, time-bin). This structure preserves the efficiency of Dijkstra's algorithm (Dijkstra 1959) while representing temporal variability (Zeng et al. 2009).

- As such, each state stores two variables: cumulative exposure and arrival time.
- The priority queue is ordered lexicographically: states with lower exposure are always expanded first, and ties are resolved by choosing the state with the earlier arrival time (Friesz et al. 2011; Zhang et al. 2012). This ensures a deterministic and unambiguous progression through the network during the pathfinding process.
- When examining an edge  $u \rightarrow v$ , the algorithm predicts when the traveler would enter and exit that edge, computes its exposure cost at that time, and updates the label at v if this new state is better than its existing record (more in Section 6.1).

Thus, the search returns the path of minimum cumulative exposure with travel time already embedded in the edge costs. Five-minute bins were chosen as an interval that offer a good balance between temporal detail and computational efficiency.

## 5.4. Outputs

For every origin–destination pair and scenario, the routing engine outputs:

- path geometry (list of nodes and edges)

- total exposure<sup>1</sup>
- travel time
- diagnostic statistics (for example, degrees-minutes per trip, exposure savings compared to shortest path, etc.)

These results serve as inputs to downstream analyses that compare baseline and mitigated scenarios or different hazard resolutions.

## 6. TIME-DEPENDENT DIJKSTRA FOR SINGLE AND MULTI-HAZARD EXPOSURE

### 6.1. Dynamic vs. Static (Snapshot) Routing

Earlier version of Icarus relied on a static approach in which every potential edge's  $T_{MRT}$  value was read at the departure time and kept fixed for the entire trip. While this method was computationally efficient, it did not represent the changing hazard conditions encountered by travelers as they moved through the network, finishing their journey. As a result, estimated exposures reflected a single momentary estimate of temperature on the links rather than the actual sequence of heat conditions experienced along the route. For instance, consider a trip that starts at 4:00 PM and takes approximately 30 minutes to finish having some candidate links on a tree-lined path and the others along the curb-front of a high-rise building. The tree-lined path may appear to be the most viable option owing to the high reflective heat contributed by the high-rise building at 4:00 PM but at 4:30 PM, the angle of the sun might result in shade being formed over the curb-front due to the presence of the high-rise building making such candidate links more desirable (Levenson et al. 2025; Shabrang et al. 2024; Zhang et al. 2024). This change in desirability of links will potentially not be captured by a snapshot of link-costs calculated at the beginning of the trip.

Icarus 25 replaces this static methodology with a dynamic, time-dependent routing framework. Operating at a 5-minute time interval, the temperature cost of each link is calculated at the precise time the traveler is expected to reach that link, enabling each downstream segment to reflect its true, time-varying microclimatic condition given the trip travel time is greater than the temporal resolution of 5 minutes. For trips with travel time shorter than this resolution, it is assumed that hazard conditions will not change significantly (see section 6.1a). If computational power allows for it, this resolution can be further improved. This allows the model to simulate traveler decisions under conditions that potentially evolve during the course of a single trip.

For a single hazard such as heat, the model uses a time-dependent Dijkstra algorithm to minimize cumulative exposure. Exposure assessment and rerouting frameworks are also available for multiple hazards (see Section 6.2).

#### 6.1a Temporal resolution and aggregation

---

<sup>1</sup> The algorithm can be updated to both account for or/and report the average exposure instead of the total exposure

To keep the time-dependent search manageable at the metropolitan scale ( $\approx 10^6$  trips), Icarus 25 discretizes time into fixed bins and prices each edge using the time bin corresponding to its predicted entry time. This entry time can be different for the same link depending on when it features in one of the several prospective paths considered by the routing algorithm. Using larger time-interval bins reduces the size of the state space and speeds computation, but it can also distort both route choice and exposure estimates because the true temperature at the moment a link is approached may not be captured precisely.

With spatial resolution held constant, we tested time steps of 1, 5, 10, and 15 minutes. Using 1-minute interpolation as the reference, we found that 5-minute bins closely matched both the benchmark routes and their associated exposure for almost all trips. By contrast, 10- and 15-minute bins produced noticeable differences. Not more than a few trips found a cooler detour when rerouting should have occurred, and some cool detours, though indeed cooler than the shortest-path baseline, did not match the benchmark-optimal path. These discrepancies stem from coarser aggregation washing out short-lived, minute-scale temperature variations along the same links.

## 6.2. Multi-Hazard Extension - Minimizing Multiple Simultaneous Hazards

Icarus 25 supports the following three methods for balancing multiple possibly competing or complementing environmental objectives.

1. **Weighted Sum Method** - Provides a fast and straightforward method to combine hazards into a single score. This method is equivalent to a single objective minimization routing in terms of computational power requirements.
2.  **$\epsilon$ -Constraint approach** - One hazard is minimized while the other is kept within a user-defined tolerance of its best possible value. This provides direct control over acceptable trade-offs and aligns better to model or inform policy restrictions.
3. **Full Pareto Frontier method** - This delivers the most comprehensive picture of trade-offs between the two hazards (for example, heat and air quality). It is the most informative but also the most computationally demanding approach.

### 6.2a Weighted Sum Formulation

The weighted sum formulation combines multiple objectives into a single composite cost:

$$C = W_H H + W_P P$$

where H and P represent cumulative exposure to heat and particulate matter respectively, and  $W_H$ ,  $W_P$  are user-specified weights reflecting their relative importance (as such, they need not sum up to 100%).

This method is simple to implement, as it allows multi-hazard routing to reuse the single-objective Dijkstra framework. However, the outcomes from this process presents significant methodological limitations:

## 1. Inability to capture non-convex trade-offs

In Figure 2, each feasible route is shown as a point on a two-dimensional plot, where one axis reflects heat exposure and the other reflects pollution exposure. Together, these points form a frontier that denotes how the two hazards trade off against one another. As we see in this example, this frontier of feasible routes can often be an irregular curve.

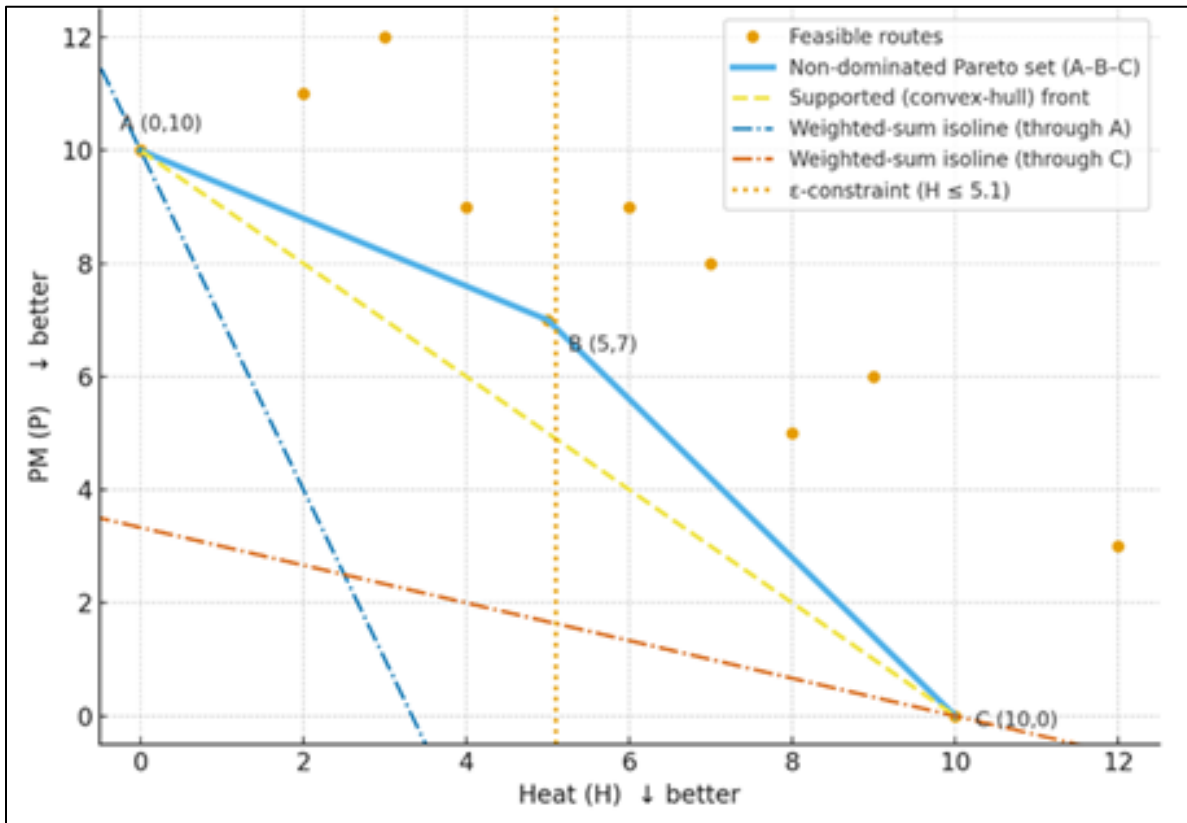


Figure 2: *Weighted-sum vs.  $\epsilon$ -constraint vs. full Pareto Optimization on a discrete, feasible set of paths*

The weighted-sum method can only find points that lie on straight sections of this 'non-convex' curve, missing the routes that lie in the dips and bends of the frontier (called "non-convex regions" of the curve). These missed routes may be the ones that planners are looking for, for instance, a path that is only 5% warmer but 40% cleaner.

## 2. Scale and Normalization Sensitivity

In a weighted-sum formulation, each hazard layer (e.g., heat,  $PM_{10}$ ) must first be converted into comparable numerical scales before being combined. However, because heat and air quality are measured in different units and vary across very different numerical ranges, the resulting composite cost is highly sensitive to how these values are scaled or normalized. A small change in the normalization window, such as whether temperatures are normalized over 35–55 °C or 25–55 °C, can significantly alter the relative importance of heat versus pollution in the optimization. The "optimal" path can then jump from one route to another, not due to any real shift in heat or pollution,

but solely due to the nature of the rescaling of the data. This causes weighted-sum results to vary across time, geography, and datasets, undermining their consistency and their value for policy-relevant interpretation.

### 3. Lack of policy control

Weighted sums cannot incorporate explicit bounds or guarantees, such as “stay within 5% of the heat-optimal route”, offering limited interpretability for policy or planning applications.

Moreover, seldom can one apply constant weights across a large study area. For pedestrian routes especially, spatial gradients of temperature and  $PM_{10}$  are expected to vary sharply across microclimates, tree or building shades, and times of day. This would indicate that a single pair of weights for the two hazards would potentially not be representative of exposure trade-offs consistently throughout the region. If this approach is used, care must be taken to assign different weights to different spatial pockets in the study area; this would include classifying microclimatic conditions, calibrate local trade-off ratios, and continuously updating them under varying temporal and meteorological contexts for the set spatial pockets. Ideally, normalization scales should be updated in a similar way, where varying gradients of temperature or PM distribution over a day would require different normalization factors over these spatial pockets. Such region-specific tuning substantially increases modeling complexity and undermines the simplicity that initially motivates the weighted-sum method.

## 6.2b $\epsilon$ -Constraint Formulation

The  $\epsilon$ -constraint approach offers a more deliberate way to balance multiple hazards by placing an explicit limit on how far a solution may drift from the best possible outcome for the primary hazard, minimizing the secondary hazard within an  $\epsilon$ -tolerance of that budget.

1. **Primary Optimization** - The first run identifies the route with the lowest exposure for the primary hazard variable (for instance, heat) and records the optimal cost  $H^*$ .
2. **Secondary optimization under constraint** – A subsequent run then minimizes the secondary hazard variable (say,  $PM_{10}$  exposure), but only among routes whose heat exposure stays within  $H \leq (1 + \epsilon)H^*$

Using the same time-dependent Dijkstra engine and lexicographic queue ordering, now recorded instead by ( $PM_{10}$ , arrival\_time), the algorithm rejects paths exceeding the specified heat budget. For each (node, time-bin) pair, the algorithm keeps the option with the lowest PM exposure or, in case of a tie, the path with the lowest heat exposure. By adjusting  $\epsilon$ , planners can decide how tightly the solution must adhere to the heat-optimal route. A small  $\epsilon$  would thus enforce strict limits on heat exposure, while a larger  $\epsilon$  would permit more flexibility in exchange for reductions in exposure to pollution.

This method potentially recovers solutions that lie in the “non-convex” portions of the trade-off curve. These are regions on the (H, P) frontier that the weighted-sum method would typically miss. In Figure 2, the  $\epsilon$ -constraint search identifies the point (H=5, P=7)

because it delivers the lowest PM exposure within the heat budget allowed to the left of the  $\epsilon = 5.1$  boundary, which the weighted formulation fails to find unless region- or trip-specific weights are implemented that could, in theory, account for this section of the graph.

Computationally, it remains lightweight as it requires only two routing passes per origin–destination pair. In practice, Icarus uses the  $\epsilon$ -constraint method by default. It strikes a good balance between accuracy and computational feasibility and directly supports policy questions framed as “how much can we reduce air pollution exposure without substantially worsening heat exposure?” and provides clear interpretability, for instance, how to find routes that stay within 5% of heat-optimal option but with minimal pollution exposure”.

## 7. ADDITIONAL ROUTING ENHANCEMENTS

### 7.1. *Spatial Precision at Trip Ends: Edge Splitting & Temporary Nodes*

Accurate routing and fair exposure estimates depend on starting and ending trips at the actual curbside or parcel locations. Snapping to the nearest network node would assign the full exposure cost of the first and/or last edge to the total trip cost even if the traveler possibly only walked part of it. This could bias both time and exposure estimates, particularly for short trips. We introduce the following workflow that inserts temporary nodes directly at the true origin and destination points. As such, splitting of the first and last traversed links ensures the exposure penalty for partial edges scales precisely with the traveled fraction.

#### **Edge Splitting Workflow**

- Nearest Edge Query - Temporarily treats the directed graph as undirected and searches for the edge whose geometry lies closest to the given point. That edge becomes the candidate for splitting.
- Projection & Validation - Computes the distance along the line (`proj_d`) where the point would fall if projected onto the edge’s geometry. If the projection lies almost exactly at either endpoint, we skip the split to avoid creating fragments that are effectively zero-length.
- Insert Temporary Node - The original edge (`u, v, key`) is removed (along with its reverse direction if the edge is bidirectional) and a temporary node (e.g., `-1` for origins, `-2` for destinations) is inserted at the projected coordinates. This is the effectively the point at which the link is split.
- Explode Edge - The original link is divided into two new geometries, i.e. the segment from `u` to the temporary node, and the segment from the temporary node to `v`. Each new edge carries the original (`u, v, key`) identifiers so cost calculations still know which underlying street segment they belong to. The upstream sub-link becomes (`u → temp_node_id`) and the downstream sub-link becomes (`temp_node_id → v`). For bidirectional links, the upstream sub-link becomes (`temp_node_id → u`) and the downstream sub-link becomes (`v → temp_node_id`). This is important as prospective paths can choose to travel this split link in either direction to achieve the minimum cost path.

Attributes other than geometry and length are preserved for the new edges, so that downstream cost routines remain functional. This is important because the cost functions rely on the original, unsplit geometries in the base walk network. By retaining both the original identity and the new split geometries, Icarus can compute exposure and travel time on the exact portion of the street the traveler actually walks, without disrupting the rest of the routing pipeline.

### *7.2. Guarding against backtracks to split points*

After splitting, both origin and destination nodes exist exactly on the network, and Icarus 25's time dependent Pareto-Dijkstra can legitimately explore the temporary split nodes. However, this also enables trivial back and forth loops, e.g., the algorithm might step from node A into the newly inserted split node -1 and then immediately back to A, because the cost on that very short half edge may appear competitive if no other candidate paths come close to the optimum cost.

Before, Icarus snapped origins/destinations to an existing node, which meant no temporary nodes and thus no opportunity for immediate reversals. To prevent these artificial back-and-forth steps without changing the underlying search, we add a simple post-processing step to the raw node sequence. As we walk through the nodes, we check whether the current node matches the one two positions earlier. When this happens, we know the path took the form  $A \rightarrow B \rightarrow A$ . We drop the intermediate node B and avoid re-adding the repeated A, so the cleaned path ends at the original first instance of A rather than traverse the unnecessary loop. Keeping this logic in a lightweight cleanup function preserves the guarantees of the main Pareto-Dijkstra search while removing any present backtracking that edge-splitting can introduce at trip endpoints.

### *7.3. Localized Subgraphs & Spatial Indexing*

Routing over a 30,000 km pedestrian network for close to a million trips is computationally expensive. In many cases, a traveler's origin and destination lie within a small corridor, so exploring the entire city graph is wasteful. To address this, a node-level R tree is built, indexing each node's spatial coordinates. This allows fast intersection queries to retrieve only the nodes (and thus edges) within a user-defined buffered envelope (or a "bounding box") around a trip's origin-destination line. The R-tree is queried with these bounding box coordinates to get candidate node IDs. We extract only this relevant portion of the network and save this as a temporary copy on which the actual routing search is performed. By limiting the search to a small local subgraph, usually only a few hundred nodes and edges, the shortest-path and time-dependent searches run far more quickly, often by an order of magnitude compared with using the full network. This also cuts the memory required for each process, which makes it possible to handle larger batches without running into resource limits.

## REFERENCES

- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1, 269–271.
- Ebi, K. L., Capon, A., Berry, P., Broderick, C., de Dear, R., Havenith, G., ... & Jay, O. (2021). Hot weather and heat extremes: health risks. *The Lancet*, 398(10301), 698-708.
- Friesz, T. L., Kim, T., Kwon, C., & Rigdon, M. A. (2011). Approximate network loading and dual-time-scale dynamic user equilibrium. *Transportation Research Part B: Methodological*, 45(1), 176-207.
- Gunawardena, K. R., Wells, M. J., & Kershaw, T. (2017). Utilising green and bluespace to mitigate urban heat island intensity. *Science of the total environment*, 584, 1040-1055.
- Guzman-Echavarria, G., Middel, A., & Vanos, J. (2023). Beyond heat exposure—new methods to quantify and link personal heat exposure, stress, and strain in diverse populations and climates: The journal temperature toolbox. *Temperature*, 10(3), 358-378.
- IPCC AR6, Working Group I, Summary for Policymakers (2021)
- Levenson, M., Pearlmuter, D., & Aleksandrowicz, O. (2025). An observational analysis of shade-related pedestrian activity. *Buildings & Cities*, 6(1).
- Li, R., Vanos, J. K., Chester, M. V., Middel, A., & Hernandez-cortes, D. (2024). Characterizing Individual Occupation and Travel Heat Burdens through Exposure and Reprieve.
- Li, R., Chester, M. V., Hondula, D. M., Middel, A., Vanos, J. K., & Watkins, L. (2023). Repurposing mesoscale traffic models for insights into traveler heat exposure. *Transportation research part D: transport and environment*, 114, 103548.
- Li, R., Chester, M. V., Middel, A., Vanos, J. K., Hernandez-Cortes, D., Buo, I., & Hondula, D. M. (2023). Effectiveness of travel behavior and infrastructure change to mitigate heat exposure. *Frontiers in Sustainable Cities*, 5, 1129388.
- Maricopa County Department of Public Health. 2021 Heat-Associated Deaths Final Report
- Middel, A., AlKhaled, S., Schneider, F. A., Hagen, B., & Coseo, P. (2021). 50 grades of shade. *Bulletin of the American Meteorological Society*, 102(9), E1805-E1820.
- Middel, A., & Krayenhoff, E. S. (2019). Micrometeorological determinants of pedestrian thermal exposure. *Science of the Total Environment*, 687, 137–146.
- Middel, A., Selover, N., Hagen, B., & Chhetri, N. (2016). Impact of shade on outdoor thermal comfort—a seasonal field study in Tempe, Arizona. *International journal of biometeorology*, 60(12), 1849-1861.
- National Weather Service Phoenix (NOAA). 2022 Climate Year in Review for Phoenix, Yuma, and El Centro. Graphic: Number of 90 degree or warmer days in Phoenix. Accessed Oct 2025.
- Shabrang, A., Heris, M. P., & Flohr, T. (2025). Measuring the Shade Coverage of Trees and Buildings in Cambridge, Massachusetts. Findings.
- Zhang, H. M., Nie, Y., & Qian, Z. (2008). Estimating time-dependent freeway origin-destination demands with different data coverage: sensitivity analysis. *Transportation research record*, 2047(1), 91-99.

Zhang, X., & Ludwig, F. (2025). Shade for pedestrians: A novel approach to calculate the spatio-temporal shade benefits of street trees considering pedestrian flow. *Building and Environment*, 272, 112662.

Zeng, W., & Church, R. L. (2009). Finding shortest paths on real road networks: the case for A\*. *International journal of geographical information science*, 23(4), 531-543.