

Capturing trade-offs between daily scheduling choices

Janody Pougala^{a,*}, Tim Hillel^b, Michel Bierlaire^a

^a École Polytechnique Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Transport and Mobility Laboratory, Switzerland

^b University College London (UCL), Department of Civil, Environmental and Geomatic Engineering, United Kingdom

ARTICLE INFO

Keywords:

Activity-based modelling
Daily scheduling behaviour
Mobility
Mixed-integer optimisation
Random utility maximisation

ABSTRACT

We propose a new modelling approach for daily activity scheduling which integrates the different daily scheduling choice dimensions (activity participation, location, schedule, duration and transportation mode) into a single optimisation problem. The fundamental behavioural principle behind our approach is that individuals schedule their day to maximise their overall derived utility from the activities they complete, according to their individual needs, constraints, and preferences. By combining multiple choices into a single optimisation problem, our framework is able to capture the complex trade-offs between scheduling decisions for multiple activities. These trade-offs could include how spending longer in one activity will reduce the time-availability for other activities or how the order of activities determines the travel-times. The implemented framework takes as input a set of considered activities, with associated locations and travel modes, and uses these to produce empirical distributions of individual schedules from which different daily schedules can be drawn. The model is illustrated using historic trip diary data from the Swiss Mobility and Transport Microcensus. The results demonstrate the ability of the proposed framework to generate complex and realistic distributions of starting time and duration for different activities within the tight time constraints. The generated schedules are then compared to the aggregate distributions from the historical data to demonstrate the feasibility and flexibility of our approach.

1. Introduction

The scheduling of daily activities is a complex process that combines multiple choices, including deciding which activities to perform in a day, as well as the timings, location, and mode-of-travel for each performed activity. These choices are not made independently. Instead, peoples' realised schedules are the end result of a series of interconnected, unobserved (and possibly unconscious) dynamics, reasoning, and trade-offs. For example, an individual might leave work earlier than usual on days they need to pick up their children from school or skip a regular exercise session entirely due to a high-workload. Being "in a rush", having "plenty of time" or being able to "squeeze in" additional activities in otherwise packed schedules are universal experiences that illustrate the trade-offs we evaluate when scheduling our days.

The daily scheduling process is a critical component of Activity-Based Models (ABMs) of transport demand, which assume that demand for transportation can be derived from the needs of individuals to perform activities (Bowman and Ben-Akiva, 2001) and that this need is influenced by space and time constraints (Chapin, 1974; Hägerstrand, 1970).

* Corresponding author.

E-mail address: janody.pougala@epfl.ch (J. Pougala).

There can largely be considered two major modelling paradigms within ABMs: rule-based and econometric models. Rule-based, or computational process, models (e.g. Golledge et al., 1994; Timmermans, 2003; Arentze and Timmermans, 2000) use decision rules to derive feasible solutions. This makes them easier to implement in practice, but the rules are hard-coded and often arbitrary, which limits their generalisation. On the other hand, econometric models postulate that scheduling can be explained with econometric processes such as random utility maximisation. As such, econometric ABMs do not typically model behaviour explicitly but rather consider it a consequence of the maximisation of utility. The different choice dimensions in an econometric model are often modelled sequentially (e.g. Adler and Ben-Akiva, 1979; Bowman and Ben-Akiva, 2001; Recker et al., 1986; Hilgert et al., 2017; Bradley and Bowman, 2008), where each decision is modelled as dependent on all previous decisions in the sequence. The decision order is decided by the modeller and therefore may not reflect that of the decision-maker, which may be recursive and not sequential. Other econometric models solve this issue by considering some or all of the choice dimensions jointly (e.g. Ettema et al., 2007, Nurul Habib, 2018; Charypar and Nagel, 2005), but a full integration of trade-offs between these choices has not yet been achieved.

In this paper, we introduce a new approach to modelling individual activity scheduling, based on mixed-integer optimisation. The key advantage of our approach is that the different modelling dimensions (activity participation, activity location, activity schedule, activity duration, and transportation mode choice to travel to the next activity) are considered jointly in a single optimisation problem. This allows the framework to capture the trade-offs that individuals evaluate when scheduling their daily activities. These trade-offs could include changing the duration of an activity to leave more time for others, choosing a specific location for an activity so as to minimise travel time, or prioritising certain activities over others. Furthermore, our approach can be used to generate an empirical distribution of individual schedules, from which different instances of daily schedules can be drawn stochastically, in order to be used in simulation. Finally, the framework is built on first behavioural principles of random utility theory and can be generalised to complex mobility situations.

Our framework focuses explicitly on activity scheduling and travel planning and so does not cover other stages typically present in operational models such as activity generation and dynamic planning (e.g. schedule updates due to unplanned events). To model the trade-offs occurring during the scheduling process, we define a mixed integer optimisation problem subject to time and cost budget constraints. We integrate explicitly the following choice dimensions:

- activity participation,
- timings (i.e. start time and duration),
- activity sequencing,
- location,
- and mode of transportation.

These choices are each subject to their own set of constraints and requirements (for example, choosing a mode requires its availability to the individual), but are interrelated. Specifically, we include the influence of both observable (e.g. technical constraints) and unobservable (e.g. personality) factors.

2. Scope

The simulation framework presented in this paper is intended to be integrated into a wider activity-based modelling process. Fig. 1 illustrates an activity-based pipeline, starting from a modelling component to estimate the parameters of the utility functions that are optimised within the simulation framework. The resulting schedules are used to derive indicators for the analysis of transport

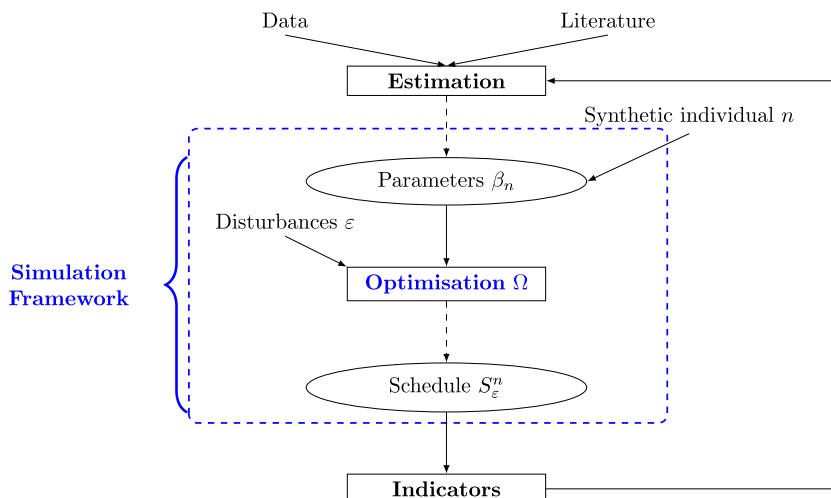


Fig. 1. Activity-based process including the simulation framework. Dashed arrows represent inputs, while solid arrows represent outputs.

behaviour, which can serve as input for agent-based microsimulators (e.g. Manser et al., 2021).

In this paper, we focus only on the simulation aspect of the problem. As such, the parameters used in the model are considered given. They can be imported from literature, or calibrated from data.

The rest of the paper is laid out as follows. Section 3 presents a brief review of the relevant literature, with an emphasis on utility-based models and simulators. The framework is detailed in Section 4, with an overview of the key components of the model and the simulation methodology. We illustrate the operation, flexibility, and realism of the framework on the Swiss Mobility and Transport Microcensus (MTMC) in Section 5. This objective of this investigation is to demonstrate that the framework is able to produce sensible results, which can later be included in an activity-based estimation of transport demand.

Finally, we conclude with a discussion on current and future challenges.

3. Relevant literature

Activity-based models originally emerged in the 1970s as a response to the shortcomings of traditional 4-step models (Vovsha et al., 2005; Castiglione et al., 2014), namely:

- trips are the unit of analysis and are assumed independent, meaning that correlations between different trips made by the same individual are not accounted for properly within the model;
- models tend to suffer from biases due to unrealistic aggregations in time, space, and within the population; and
- space and time constraints are usually not included.

The early works of Hägerstrand (1970) and Chapin (1974) established the fundamental assumption of activity-based models, that the need to do activities drives the travel demand in space and time. Consequently, mobility is modelled as a multidimensional system rather than a set of discrete observations. Unlike traditional trip-based models, ABMs focus on overall behavioural patterns: decisions are analysed at the level of the household as opposed to seemingly independent individuals, and dependencies between events are taken into account (Timmermans, 2003; Pas, 1985). Specifically, modellers are interested in the link between activities and travel, often considered within a given timeframe. Typically, a single day is used as the unit of analysis. The resulting goal of studies in the literature is therefore to replicate as accurately as possible the interactions and considerations involved in the development of a daily schedule by an individual.

While the scheduling process is central to the activity-based research, there is no clear consensus on the representation and modelling of the daily scheduling process in utility-based frameworks. Typically, individuals are assumed to schedule activities by maximising the utility they can expect to gain. The timeframe is often introduced as a time budget that constrains the overall time expenditure. The scheduling decisions can be modelled as discrete choices: sequential discrete choice models consider a series of choices done consecutively with varying amounts of feedback between each step. On the other hand, joint models also integrate correlations between each aspect of the scheduling decision by evaluating them simultaneously. Other models do not consider the choice as fully discrete, but an hybrid consumption of discrete and continuous “goods”.

The earliest functional utility-based models are sequential models such as the logit model for household daily travel patterns developed by Adler and Ben-Akiva (1979), which assumes that households choose from a set of possible daily patterns and uses a logit model to compute the choice probabilities for each alternative. It was followed by the disaggregate travel demand model developed by Bowman and Ben-Akiva (2001) that models a series of sequential decisions to generate an activity pattern and tours for the day. These decisions are:

- 1 the choice of activity pattern (staying at home or travelling),
- 2 the primary tour time of day,
- 3 the primary tour destination and mode,
- 4 the secondary tours times of day, destination and modes.

The choice of activity pattern is modelled using a nested logit model, the tour times of day are generated using a logit model, and the destination and mode with a logit model with alternative sampling. A set of rules is used to define a hierarchy among activities (primary vs. secondary). The models developed by Adler and Ben-Akiva, Bowman and Ben-Akiva are travel-centric: while both assume an interdependence among choices, they mostly focus on trip characteristics (e.g. tour frequency, number of stops, mode choice ...). Behavioural mechanisms explaining the actual choice of activities and their sequence are examined less closely. In the context of these models, activity schedules and emerging behaviour are implicit and rather consequential to the predicted travel decisions.

Sequential estimation remains popular in the literature, especially for microsimulators (e.g. Recker et al., 1986; Pendyala et al., 2005; Smith et al., 1995; Ettema et al., 2000; Axhausen et al., 2016) that have expanded its predicting possibilities. STARCHILD (Recker and Root, 1981; Recker et al., 1986), is one of the first of many operational microsimulators. STARCHILD models activity-travel decisions as a sequence of five stages (household interactions and individual activity programs, scheduling, recognition of activity patterns, specification of the choice set, and choice model for the activity pattern) in order to simulate the choice of a daily activity schedule, including planned and unplanned activities, and interactions with household members. Specifically, the scheduling model takes as input the program (set of planned activities and their spatio-temporal characteristics) and generates different combinations of this plan. The resulting set of schedules is evaluated against feasibility constraints, then used as input for a multi-objective optimisation model that simulates the choice set of an individual. Estimating the choices in a sequence allows for simple, clearly

defined modelling assumptions, but limits the ability of the framework to capture trade-offs that individuals could make between different choice dimensions. For instance, STARCHILD offers a robust solution to generate the planned set of activities from a larger set of possibilities — or the “opportunity set” (Recker et al., 1986) — but this plan cannot be revised in later stages of the scheduling process (e.g. adding/dropping activities, increasing/decreasing time). Another limitation is the deterministic aspect of the generated alternatives, which relies on full enumeration of the choice set, followed by a reduction with decision rules.

More recent works have focused on joint estimation of mobility choices, with a more explicit integration of emerging behaviour in the scheduling process. For instance, Nurul Habib and Miller (2009) use an utility-based approach to model the generation of activities (i.e. which activities are considered in the first place). In this case, the utility function is defined for an agenda (a set of activities to be scheduled), aiming to capture the trade-off between planned and unplanned activities. The choice probabilities are estimated with the Kuhn-Tucker optimality conditions in place of discrete choice models. The resulting agenda is then used as input for a discrete-continuous scheduling model that predicts sequentially the choice of activity (discrete choice) and the time expenditure for the chosen activity (continuous choice) (Nurul Habib, 2011). This theoretical framework is the foundation of CUSTOM, a utility-based scheduling model of workers’ daily activities (Nurul Habib, 2018), simulating the discrete choice of performing an out-of-home activity or staying at home all day, and in the former case, the choice of start time of the first trip. The framework goes through multiple “scheduling cycles” to model activity-travel choices such as the choice of activity type, duration, mode and location, for every activity to be scheduled — subject to a time budget. Each scheduling cycle modifies the remaining time budget, which is used to generate a Potential Path Area of feasible locations and mode available for the next activity. The authors model trade-offs between activity cycles and scheduling decisions by assuming that the expected utility of the type and location choice of the next activity impacts the utility of the current cycle. The process is therefore not fully simultaneous: the scheduling cycles are evaluated in a series, with each episode influencing and being influenced by the following. This approach provides greater flexibility than the strictly sequential models that have been presented, but limits the ability of the model to capture more complex interactions (e.g. influences of multiple activities).

The discrete-continuous representation of activity schedules has been investigated extensively by Bhat et al. (2004) (see also Bhat (2005; 2018)). In their Multiple Discrete-Continuous Extreme Value (MDCEV) model, the scheduling process is modelled as a combination of a discrete choice (activity participation) and continuous choice (activity duration). Behaviour is explicitly considered with a non-linear utility function and satiation effects (decreasing marginal utility). The discrete-continuous approaches are a flexible solution to simultaneously consider multiple choice dimensions. However, they become limited when it comes to integrate time-of-day decisions, which are heavily influenced by external factors (e.g. shop opening times, working hours, commitments, etc.). To address this issue, Palma et al. (2021) propose to modify the MDCEV formulation to estimate the durations of activity episodes instead of activity types, by considering a maximum number of episodes per activity, and including a polynomial penalty and a satiation parameter in the utility function of each activity. This approach allows to capture trade-offs between activities - but is limited to the choice of activity participation and duration, and therefore needs to be combined with a scheduling model in order to be used in an activity-based context.

Joint estimation of multiple choice dimensions, including time-of-day, has been explored in other works. Ettema et al. (2007) formulate an error-component discrete choice model to jointly estimate duration, time-of-day preference and effect of schedule delays on the utility function of the alternatives. They consider that individuals maximise the sum of the utility gained from travelling and from performing the activities, the latter composed of three elements: a time-of-day dependent utility, a duration utility, and a schedule delay utility dependent on the start time. Their model is thus able to accommodate more explicitly the discontinuities in utility introduced by the presence of these external constraints and preferences. However, it mainly focuses on time allocation for a given set of activities, and schedule dynamics linked to activity participation (e.g. dropping an activity if the timings are not convenient for the individual) cannot easily be taken into account.

Several key features appear in the reviewed methodologies and operational models. These characteristics contribute to the behavioural realism of the approaches:

1. Simultaneous estimation of choices: the scheduling choice (activity type, time expenditure, mode, location ...) are estimated jointly, which increases the ability of the model to deal with interactions and correlations;
2. Activity participation: the model includes the choice of participating to a set of possible activities, as opposed to only scheduling a pre-defined set of activities;

Table 1

Modelling features of the state of the art and practice.

Models	Features						
	Simultaneous estimation	Activity participation	Continuous time	Explicit behaviour	Social system	Resources	Trade-offs
Recker et al. (1986)		X			X	X	
Miller and Roorda (2003)					X		
Bhat et al. (2004)	X		X	X	X		
Ettema et al. (2007)	X	X	X	X			
Nurul Habib (2018)	X	X	X	X	X	X	
Palma et al. (2021)		X	X	X			X
Current research	X	X	X	X	(X)	(X)	X

3. Continuous time representation: time is modelled as continuous or with fine granularity in order to obtain schedules rich enough for a variety of applications;
4. Explicit modelling of behaviour: behavioural elements influencing the scheduling choices (e.g. preferences, flexibility, satiation, etc.) are explicitly modelled, and their effect can be easily interpreted;
5. Social system: the model includes the impact of social interactions, at the level of the household or larger circles;
6. Resource availability: resources such as mobility tools (private vehicles, public transport subscriptions, ...) or income are included in the model and impact the availability of certain alternatives to the decision-makers;
7. Scheduling trade-offs: the model is able to capture trade-offs in schedule timings (i.e. compromises on timings, for example to accommodate more or longer activities), participation to activities, location, mode and route choices.

Table 1 maps the discussed methods with these components, and highlights the contribution of our research. In the table, X is used to identify a feature that is available in the corresponding model. (X) is used for our methodology, and defines features that have not yet been implemented in the initial state of the framework as described in this paper, but are feasible extensions (as detailed in Section 4.4).

Reliable estimation and simulation of activity-travel behaviour requires a framework that includes every choice pertaining to the activity-travel behaviour (participation, scheduling, destination and mode of transportation), and is able to deal with their correlations. This procedure is individual-specific, with parameters and error terms of the utility function to be distributed across the population, but the optimisation can be performed at higher dimensions with the integration of household dynamics. The model should provide enough flexibility to accommodate different cases and constraints, and be easily used in conjunction with or as input for powerful tools such as agent-based models or traffic simulators.

The research presented in this paper lays the foundation of the framework, with the development of the core optimisation-based simulation of the scheduling process. We focus on the modelling of single-day scheduling. Several authors have pointed that this ignores day-to-day correlations and dynamics (e.g. Arentze et al., 2011). The framework can be extended to include multi-day dynamics, but at a higher computational cost.

4. Modelling framework

The framework presented in this paper captures the choice of a valid schedule for a given time horizon (typically a day, though it could be any time period) made by a single individual, called the agent. The central theory behind our approach is that individuals schedule their day to maximise their overall derived utility from the activities they complete, according to their individual needs, constraints and preferences. We therefore define a general utility function which captures the derived utility from an individual completing a considered activity. The form of the utility function is flexible, and can include (but is not limited to) features capturing the behaviour of the individual related to the given activity and the related trip(s). Additional factors influencing the scheduling process (e.g. interactions with other agents, routine effects ...) can be included in the utility function, with a specification to be defined according to the desired trade-off between model realism and computational accuracy. In the early stages of our work, and in the rest of the paper, we consider a utility function that includes the following variables:

- the preference towards *participating* in that the type of activity,
- the desired and scheduled *duration* of the activity,
- the desired and scheduled *start-time* of the activity,
- the flexibility of the individual towards *schedule deviation* in start-time (early/late) and duration (long/short) for the activity,
- the *cost* of participating in the activity, and
- the required *travel time* and *travel cost* to arrive at the activity location from the previous location.

We then define a mixed-integer optimisation problem for each individual which maximises the sum of the utilities of each completed activity in a schedule over a fixed time horizon. This optimisation problem can therefore capture the *trade-offs* between scheduling decisions for multiple activities, such as how spending longer in one activity will reduce the time-availability for other activities or how the order of activities changes the travel-times. The overall framework takes as input a set of *considered activities*, with associated locations and travel modes, and uses this to define a distribution over possible schedules, from which likely scheduling choices can be stochastically drawn.

In this section, we introduce the modelling elements of the proposed framework. For the sake of clarity of the notations, no index is associated with the agent in the following analysis.

Time can be either continuous or discrete. The time horizon starts at $t = 0$ and finishes at $t = T$. Space is characterised by a discrete and finite list of L locations, indexed by ℓ . The location $\ell = 0$ is called “home”, and is assumed to be the location of the agent at time $t = 0$ and time $t = T$. As well as a fixed time horizon, the agent is assumed to have a maximum daily budget B to cover the costs of activity participation and travel.

The agent considers M transportation modes, indexed by m . The travel time between two locations ℓ_o and ℓ_d using mode m is denoted $\rho(\ell_o, \ell_d, m)$ and is exogenous. If ℓ_d cannot be reached from ℓ_o using mode m , then $\rho(\ell_o, \ell_d, m) = +\infty$. Similarly, the travel cost between locations, which is also exogenous, is denoted as $\kappa(\ell_o, \ell_d, m)$.

The agent considers a set of A activities, indexed by a . Each activity a is associated with:

- a list L_a of possible locations where the activity could be performed,
- an indicator μ_a that is 1 if the activity is mandatory and 0 if it is optional,
- a time interval when the agent prefers to start the activity¹: $[x_a^-, x_a^+]$, where $x_a^- \leq x_a^+$,
- a minimum duration τ_a^{\min} ,
- a range of desired durations $[\tau_a^-, \tau_a^+]$, where $\tau_a^{\min} \leq \tau_a^- \leq \tau_a^+$,
- a cost c_a for participating in the activity.

Each relevant pair activity/location is associated with a feasible time interval $[\gamma_{a\ell}^-, \gamma_{a\ell}^+]$. It stipulates that the activity can take place only during that time interval. For example, shopping can typically only happen during the opening hours of the selected shop. Note that the agent may consider a location for an activity even if there is no overlap between $[\gamma_{a\ell}^-, \gamma_{a\ell}^+]$ and $[x_a^-, x_a^+]$. While the former represents a hard constraint, the latter represents a preference.

4.1. Valid schedules

Given the above information, the agent considers valid schedules. A schedule is the outcome of the agent’s decisions with respect to activity participation, activity location, activity scheduling, and transportation mode choice. More specifically, a schedule \mathcal{S} is a sequence of n activities $A = (a_0, \dots, a_n)$, starting with a dummy activity a_0 called “dawn”, and finishing with a dummy activity a_n called “dusk”, both of which take place at home. Each activity a is associated with an actual location ℓ_a , an actual starting time x_a and an actual duration τ_a . With the exception of the last activity “dusk”, a trip is performed immediately after each scheduled activity a , using an actual mode of transportation m_a . Note that, if the next activity takes place at the same location, the duration of the trip is simply zero.

A schedule is *valid* if

- it spans the whole time horizon, that is if

$$\tau_{\text{dawn}} + \tau_{\text{dusk}} + \sum_{n=1}^{N-1} (\tau_{a_n} + \rho(\ell_{a_n}, \ell_{a_{n+1}}, m_{a_n})) = T, \tag{1}$$

- it does not exceed the maximum budget, that is if

$$\sum_{n=1}^{N-1} (c_{a_n} + \kappa(\ell_{a_n}, \ell_{a_{n+1}}, m_{a_n})) \leq B, \tag{2}$$

- each activity starts when the trip following the previous activity is finished, that is

$$x_{a_{n+1}} = x_{a_n} + \tau_{a_n} + \rho(\ell_{a_n}, \ell_{a_{n+1}}, m_{a_n}), \forall n = 0, \dots, N - 1, \tag{3}$$

- the duration of each activity is valid, that is if

$$\tau_{a_n} \geq \tau_a^{\min}. \tag{4}$$

- all mandatory activities are included,
- only one activity from a set of considered duplicates (i.e. same activity with different associated locations or modes) is included in the schedule.

Further constraints on valid schedules can be included to ensure consistent behaviour. Regarding mode choice, we constrain the choice of mode of travel from the location of activity a to activity b to a set of feasible modes given the previous mode choices. For example, if a traveller takes public transport to go to work in the morning, her private car is no longer available to be chosen for other trips, until she returns home when her car will be available again. This behaviour can be generalised to all private vehicles.

4.2. Preferences

The agent is assumed to be rational, and to select the preferred schedule among all possible valid schedules. The preferences of the agent are captured by a utility function $U_{\mathcal{S}}$ associated with each schedule \mathcal{S} .

From the point of view of the analyst, the main challenge is that the choice set cannot be enumerated, due to the combinatorial structure of the set of valid schedules. We propose to address this challenge by performing an explicit enumeration at the activity level

¹ Note that the assumption that preferences in starting time and duration are captured by a unique time interval is mathematically convenient, but may not be realistic. For instance, a student may prefer to sit an exam either early in the morning, or late in the afternoon. In that case, it would be modelled using two different activities.

for decisions related to activity location and transportation mode; and an implicit enumeration for decisions related to activity participation and activity scheduling.

For each activity considered by the agent, we explicitly enumerate all possible combinations of the locations and modes associated with it. Each of these combinations is considered as a separate activity in the model. Therefore, each activity a considered by the agent is modelled by the analyst using up to ML_a mutually exclusive activities, each associated with a unique location ℓ_a and a unique mode of transportation m_a . In addition, we impose the constraint that at most one of these duplicate activities can be selected in a given schedule. This explicit enumeration leads to K groups G_k of activities that are mutually exclusive. We can therefore simplify some notations: (i) the feasible time interval of activity a can be denoted $[\gamma_a^-, \gamma_a^+]$, (ii) the travel time between two activities can be denoted

$$\rho_{ab} = \rho(\ell_a, \ell_b, m_a), \tag{5}$$

and similarly (iii) the travel cost between two activities can be denoted

$$\kappa_{ab} = \kappa(\ell_a, \ell_b, m_a). \tag{6}$$

The implicit enumeration consists of solving the scheduling problem considered by the agent using a standard optimisation algorithm, that identifies the optimal solution without complete enumeration.

Before describing the scheduling problem, we introduce the model of the utility $U_{\mathcal{J}}$ associated by the agent with the schedule. We define it as the sum of a generic utility U associated with the whole schedule and utility components capturing the activity-travel behaviour:

$$U_{\mathcal{J}} = U + \sum_{a=0}^{A-1} \left(U_a^1 + U_a^2 + U_a^3 + \sum_{b=0}^{A-1} (U_{a,b}^4 + U_{a,b}^5) \right). \tag{7}$$

The components and the associated assumptions are defined as follows:

1. A generic utility U that captures aspects of the schedule that are not associated with any activity. For instance, the agent may prefer that all shopping activities take place in the afternoon, or may dislike days with too many activities.
2. The utility U_a^1 associated with the participation of the activity a , irrespective of its starting time and duration. This term may include any variable such as level of service, cost, etc. Here, we illustrate the framework with a specification involving cost. It may also include an error term, capturing the unobserved variables.

$$U_a^1 = \beta_{\text{cost}} * c_a + \epsilon_1 \tag{8}$$

3. the utility U_a^2 associated with starting time. This term captures the perceived penalty created by deviations from the preferred starting time. Here, we illustrate this using a deterministic (dis)utility:

$$U_a^2 = V_a^2 \tag{9}$$

with:

$$V_a^2 = \theta_{a_k}^- \max(0, x_a^- - x_a) + \theta_{a_k}^+ \max(0, x_a - x_a^+), \tag{10}$$

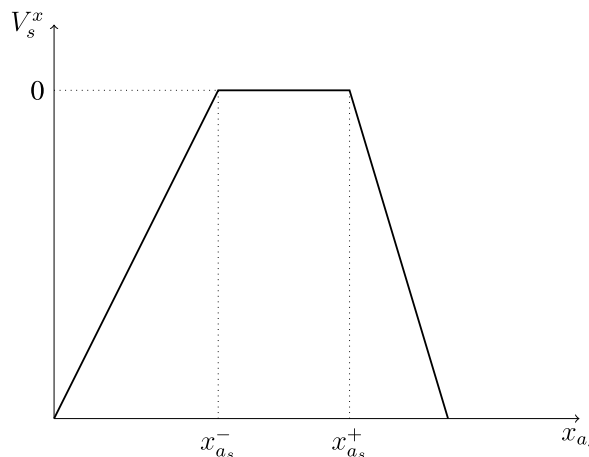


Fig. 2. Utility associated with deviations from the preferred starting time of an activity.

where $\theta_{a_k}^e \leq 0$ and $\theta_{a_k}^l \leq 0$ are unknown parameters to be estimated from data. The first (resp. second) term captures the disutility of starting the activity earlier (resp. later) than the preferred starting time, as illustrated in Fig. 2. Note that the amplitude of the penalty, captured by the parameters θ , may vary across groups of activities. The index k captures the level of flexibility with respect to the scheduling of the activity.

4. the utility U_a^3 associated with duration. This term captures the perceived penalty created by deviations from the preferred duration. Here, we illustrate this using a deterministic (dis)utility:

$$U_a^3 = V_a^3 \tag{11}$$

with:

$$V_a^3 = \beta_{a,k}^e \max(0, \tau_{a_k}^- - \tau_a) + \beta_{a,k}^l \max(0, \tau_a - \tau_{a_k}^+), \tag{12}$$

where $\beta_{a_k}^e \leq 0$ and $\beta_{a_k}^l \leq 0$ are unknown parameters to be estimated from data. Similarly to the specification of start time, the first (resp. second) term captures the disutility of performing the activity for a shorter (resp. longer) duration than the preferred one,

5. For each pair of locations (ℓ_a, ℓ_b) , respectively, the locations of activities a and b with $a \neq b$, the utility $U_{a,b}^4$ associated with the trip from ℓ_a to ℓ_b , irrespective of the travel time. This term may include variables such as cost, level of service, etc. Here, we illustrate the framework with a specification involving travel cost. It may also include an error term, capturing the unobserved variables.

$$U_{a,b}^4 = \beta_{t,\text{cost}} * c_t + \epsilon_4 \tag{13}$$

6. For each pair of locations (ℓ_a, ℓ_b) , the utility $U_{a,b}^5$, which captures the penalty associated with the travel time from ℓ_a to ℓ_b . Here, it is assumed to be deterministic:

$$U_{a,b}^5 = V_{a,b}^5 \tag{14}$$

with

$$V_{a,b}^5 = \theta_t \rho_{ab}, \tag{15}$$

where θ_t is an unknown parameter to be estimated from data, and ρ_{ab} is the travel time to the next location.

Note that no utility is associated with the dummy activity “dusk”. We also normalise

$$U_0^1 = U_0^2 = U_0^3 = 0. \tag{16}$$

Indeed, as only differences of utility matter, the two dummy activities serve as reference and their utility is set to zero.

4.3. Schedule optimisation

As is consistent with random utility theory, the agent is assumed to select the valid schedule with the highest utility. She therefore solves an optimisation problem to maximise the utility function under the validity constraints. However, from the point of view of the analyst, the utility function (7) is captured by a random variable, and the model associates a choice probability with each valid schedule. In order to deal with this uncertainty, we propose a simulation approach, where the optimisation problem is explicitly solved for several realisations of the random utility. This is done by drawing from the distributions of the random terms (e.g. ϵ_1 and ϵ_4 in the specification illustrated above), and solving the optimisation problem. Assuming a normal distribution for these quantities makes the sampling more convenient, but is not required by the framework. The resulting schedule is a realisation from the choice model. The advantage of this approach is that each generated schedule is valid by design, explicitly capturing the trade-offs made by the agent.

For each activity a , we first generate realisations of U^1 , that we denote by V_a^1 . For each pair (ℓ_a, ℓ_b) of locations, we also generate realisation of U^4 , that we denote by V_{ab}^4 . We characterise the decision of the agent using the following decision variables:

- ω_a : binary variable that is 1 if activity a is selected in the schedule, and 0 otherwise,
- z_{ab} : binary variable that is 1 if activity b is scheduled immediately after activity a , where $a \neq b$,
- x_a : starting time of activity a ,
- τ_a : duration of activity a ,
- α_a^m : indicator variable that is 1 if private mode m (e.g. car or bicycle) is available for activity a , and 0 otherwise.

We denote the corresponding vectors by $\omega, z, x, \tau, \alpha^m$. We consider a realisation of the generic utility U , denoted by $U(\omega, z, x, \tau, \epsilon)$, to emphasise that it depends on the decision variables and on the error terms.

The objective function is derived from (7):

$$\max_{\omega, z, x, \tau} U(\omega, z, x, \tau, \epsilon) + \sum_{a=0}^A \omega_a (V_a^1 + V_a^2 + V_a^3) + \sum_{a=0}^A \sum_{b=0}^A z_{ab} (V_{ab}^4 + V_{ab}^5). \quad (17)$$

The constraints are

$$\sum_a \sum_b (\omega_a \tau_a + z_{ab} \rho_{ab}) = T, \quad (18)$$

$$\sum_a \sum_b (\omega_a c_a + z_{ab} K_{ab}) \leq B, \quad (19)$$

$$\omega_{\text{dawn}} = \omega_{\text{dusk}} = 1, \quad (20)$$

$$\tau_a \geq \omega_a \tau_a^{\min}, \quad \forall a \in A, \quad (21)$$

$$\tau_a \leq \omega_a T, \quad \forall a \in A, \quad (22)$$

$$z_{ab} + z_{ba} \leq 1, \quad \forall a, b \in A, a \neq b, \quad (23)$$

$$z_{a, \text{dawn}} = z_{\text{dusk}, a} = 0, \quad \forall a \in A, \quad (24)$$

$$\sum_a z_{ab} = \omega_b, \quad \forall b \in A, b \neq \text{dawn}, \quad (25)$$

$$\sum_b z_{ab} = \omega_a, \quad \forall a \in A, a \neq \text{dusk}, \quad (26)$$

$$(z_{ab} - 1)T \leq x_a + \tau_a + z_{ab} \rho_{ab} - x_b, \quad \forall a, b \in A, a \neq b, \quad (27)$$

$$(1 - z_{ab})T \geq x_a + \tau_a + z_{ab} \rho_{ab} - x_b, \quad \forall a, b \in A, a \neq b, \quad (28)$$

$$\sum_{a \in G_k} \omega_a \leq 1 \quad k = 1, \dots, K, \quad (29)$$

$$\alpha_a^m = 1 \quad \forall a \in G_{\text{home}} \quad (30)$$

$$\omega_a \leq \alpha_a^m \quad \forall a \in A^m \quad (31)$$

$$\omega_a \geq \omega_b + z_{ab} - 1 \quad \forall a \in A, b \in A \setminus G_{\text{home}} \quad (32)$$

$$\omega_b \geq \omega_a + z_{ab} - 1 \quad \forall a \in A, b \in A \setminus G_{\text{home}} \quad (33)$$

$$x_a \geq \gamma_a^-, \quad \forall a \in A, \quad (34)$$

$$x_a + \tau_a \leq \gamma_a^+, \quad \forall a \in A. \quad (35)$$

Equation (18) constrains the total time assigned to the activities in the schedule (sums of durations and travel times) to be equal to the time horizon. Similarly, equation (19) constrains the total cost of the schedule (sums of the costs of participating and travelling to the activities in the schedule) to not exceed the maximum budget. Equation (20) ensures that each schedule begins and ends with the dummy activities *dawn* and *dusk*. Equations (21) and (22) enforce consistency with the activity duration by requiring the activity to have a duration greater or equal than the minimal duration (4) and for the activity to have zero duration if it does not take place. Equations 23–27 constrain the sequence of the activities: (23) ensures that two activities *a* and *b* can only follow each other once (thus can only be scheduled once). As it is defined for distinct activities only, it also ensures that an activity cannot follow itself. Equations 24–26 state that each activity has only one predecessor (excluding the first activity), and each activity only one successor (excluding the last activity). Equation (27) enforces time consistency between two consecutive activities (with travel time ρ_{ab}). Equation (29) ensures that only one activity within a group of duplicates *G* is selected. Equations 30–33 define the constraints related to the choice of mode of transportation. (30) ensures that all private modes *m* are always available for activities (or trips) starting from home.² (31) only allows alternatives associated with a private mode *m* to take place if *m* is available, while (32) and (33) enforce mode consistency between two consecutive activities, excluding returns home where a different (private) mode can be chosen. Finally, (34) and (35) are time-window constraints.

² This assumption can be relaxed to take into account that household dynamics influence the share of privately owned vehicles.

Note that all the constraints in this formulation are linear in the decision variables. Therefore, if the objective function is also expressed as a linear function of the constraints, we obtain a linear integer optimisation problem,³ that can be solved by standard mathematical programming algorithms (e.g. branch-and-bound, branch-and-cut, constraint programming, etc.).

Hence, we add the assumption that the generic utility $U(\omega, z, x, \tau, \epsilon)$ must be specified as a linear function of the decision variables. This assumption is common in the mathematical programming literature, and is not overly restrictive.

4.4. Flexibility of the framework

The form of the utility function specified in the previous section is highly flexible, and allows for the modeller to:

- impose arbitrary behavioural assumptions using different constraints,
- introduce additional choice dimensions.

We have shown that some specific activity-travel behaviour can be included in the model through the constraints, such as mode consistency between consecutive alternatives, or mode changes at home. The constraints do not impact the specification of the utility functions, which allows for straightforward generalisation of the model to varied behaviours and scenarios. In the next section we introduce an application of the framework for planning of a single day, but the time horizon can be specified to be any arbitrary period (e.g. a week). For longer time periods, the inclusion of *linked activities* is important to ensure consistent behaviour. Linked activities can be activities that must be included in a strict sequence (i.e. one activity performed immediately after the other), in a certain order (i.e. so that other activities may be performed in-between), or that either all or none of these activities must be included in the schedule. It is also possible to associate different potential *home* locations to model long-distance commuters who may spend some nights out of home, for instance.

Similarly, it is straightforward to include further choice dimensions without altering the specification of the utility functions. Our framework illustrates the inclusion of mode choice, though the process can be generalised to any desirable choice addition (or restriction) by including it in the definition of a considered activity. For instance, it is straightforward to extend the framework to include route choice, by associating each activity a to a specific route r_a , and duplicating the activities as many times as there are possible route alternatives.

Clearly, these extensions come at the cost of computational burden, and the trade-off between computational time and model realism must be carefully considered.

5. Empirical investigation

In order to illustrate the optimisation-based simulation concept introduced in Section 4, we rely on a real-world dataset of historic activity schedules to generate the inputs. The objective is to show that, given sets of possible activities, locations, modes and timing preferences, the model is able to generate realisations of chosen daily schedules.

The MTMC is a Swiss nationwide survey gathering insights on the mobility behaviours of local residents ([Office fédéral de la statistique and Office fédéral du développement Territorial, 2017](#)). Respondents provide their socio-economic characteristics (e.g. age, gender, income) and those of the other members of their household. Information on their daily mobility habits and detailed records of their trips during a reference period (1 day) are also available. The 2015 edition of the MTMC contains 57'090 individuals, and 43'630 trip diaries. We use only the data corresponding to the residents of Lausanne, for a total of 2'227 diaries.

5.1. Inputs

The required inputs (activities, locations, feasible and desired start times and durations, flexibility, etc.) are not always available in traditional travel surveys, including the MTMC. The challenge is thus to provide heuristics to obtain estimators for the missing attributes.

Table 2 summarises the data requirements for the operational model, as well as two possible solutions to overcome the lack of information for each requirement. The *rigorous solution* column describes a methodology to obtain the associated information without or very little simplifying assumptions or proxies. These solutions might require a dedicated model or additional data. Therefore, we have provided the *heuristic* column, with less rigorous but easier to implement alternatives. The methods described in the *heuristic* column have been applied in this paper, with results presented in section 5.3.

5.2. Utility specification

Allowing for the available inputs for this case study, the schedule utility function expressed in (7) has been simplified as follows:

1. We assume that the random terms are randomly distributed: $\epsilon_s \sim \mathcal{N}(0, \sigma^2)$, with variance σ^2 set to 1.

³ If time is modelled using a continuous variable, we solve a linear mixed integer optimisation problem.

Table 2
Data requirements for operational model.

Requirements	Rigorous solution	Heuristic
Considered activities A	Activity choice set generation algorithm for each individual	Description of actual schedule from dataset
Considered modes M	Mode choice set generation algorithm for each individual	Consider all 5 main modes (driving, passenger, public transport, walk, cycle)
Considered locations L_a	Location set generation algorithm for each individual	Description of actual schedule from dataset
Desired start time and duration ranges $[x_a^-, x_a^+]$ and $[\tau_a^-, \tau_a^+]$,	Habit analysis and identification of typical timings in multi-day diaries	Ranges replaced by recorded values in dataset
Flexibility k	Habit analysis in multi-day diaries — flexibility would be the timing variability	Assign a discrete flexibility profile to each activity based on literature classification.
Penalty values (θ, β)	Calibrated on data — n -dependent	From literature, homogeneous across all population
Feasible time windows $[\gamma_{a\ell}^-, \gamma_{a\ell}^+]$	Data collection	Out-of-sample distributions of start and end times for each activity, across the population
Minimum duration τ_a^{\min}	Habit analysis in multi-day diaries	Set to 0

2. The ranges of start time preferences $[x_a^-, x_a^+]$ are such that $x_a^- = x_a^+ = x^*$, and the associated utility V^2 is therefore defined as:

$$V^2 = \theta_{a_k}^e \max(0, x_a^* - x_a) + \theta_{a_k}^e \max(0, x_a - x_a^*), \tag{36}$$

The same assumption is made for the preferred durations and their associated utility V^3 , similarly defined as:

$$V^3 = \beta_{a_k}^e \max(0, \tau_a^* - \tau_a) + \beta_{a_k}^e \max(0, \tau_a - \tau_a^*), \tag{37}$$

3. The flexibility in time k is modelled using a discrete indicator that can describe 3 possible behaviours (Fig. 3):
- (a) Flexible (F): deviations from preferences for activity a are relatively unimportant, thus are less penalised.
 - (b) Moderately flexible (MF): deviations from preferences are moderately undesirable, and so are penalised more than in the flexible case.
 - (c) Not flexible (NF): deviations from preferences are strongly undesirable, and are consequently highly penalised.

Each activity is associated with one level of flexibility, and each level is characterised by specific values of the penalty parameters. The flexibility assignments for each activity are summarised in Tables 3 and 4. For the sake of simplicity, we consider that the parameters are deterministic instead of randomly distributed across the population. We have chosen values based on results from the departure time choice literature (Small, 1982). Similarly, we have used cost variables (travel cost c_t and cost of activity participation c_a) and associated parameters ($\beta_{t,\text{cost}}$ and β_{cost}) from the value of time literature, and specifically case studies in Switzerland:

1. National averages of travel cost per mode and distance (BFS, 2021) were used to approximate c_t ,
2. The travel cost parameters $\beta_{t,\text{cost}}$ were derived from national averages of value of time for each mode (Weis et al., 2021),
3. The Swiss Household Budget Survey (BFS, 2007) provides average expenditures for activities as a percentage of the household budget. These were used to define c_a . For the sake of simplicity, the cost of activities that are not associated with the consumption of goods (e.g. work) was set to 0.
4. The activity cost parameters β_{cost} were derived from the average value of leisure and value of time assigned to work estimated for Zurich (Schmid et al., 2021).

5.3. Results

We present four examples from the MTMC: two students, identified as Alice and Bryan; a worker, Claire, and an unemployed person, Dylan. The set of considered activities, timing preferences, activity locations and modes for each individual are reported in Table 5. Certain activities were duplicated to offer different mode and location options. For each individual, we take separate draws of the error terms ϵ , and use them to draw different optimal schedules according to the utility specification. The schedules for each individual are shown in Figs. 4–7. Like for any simulation analysis, a large number of draws is necessary. In this example, we have generated 100 realisations of the schedules, out of which we have arbitrarily selected three of them for the sake of illustration.

For Alice, all solutions show sequences where both of the *education* instances are scheduled. Regarding the *leisure* activity, only the second schedule (Fig. 4b) includes it with timings consistent with her preferences. For the other two solutions (Fig. 4a and c), this activity is scheduled at a different time of day than the desired times (in the morning and at lunch time, respectively).

For Bryan, the first two solutions both include shopping, but at different locations. In the third solution (Fig. 5c), the shopping activity does not appear in the schedule, indicating that staying at home has a higher overall utility.

The solutions for Claire (shown in Fig. 6) are similar in that all include *work*, with timings that do not diverge substantially from the preferences. On the other hand, the discretionary activities provided as input (in this case, errands, escort, leisure and shopping) are not always scheduled, and when they are, the scheduled timings can be far from the preferences (e.g. Fig. 6c).

Dylan differs from the other selected individuals in that his set of considered activities does not contain any highly constrained

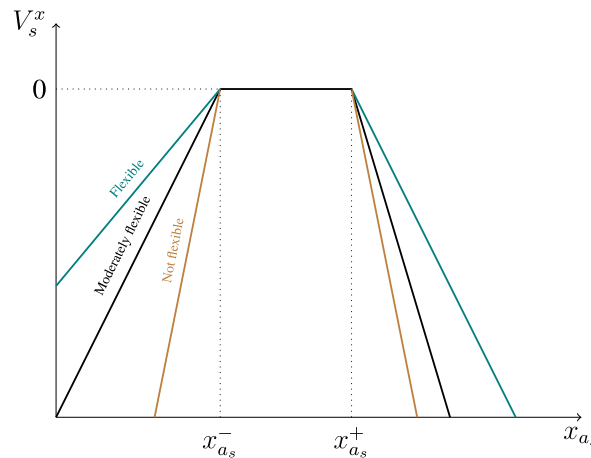


Fig. 3. Utility associated with deviations from the preferred starting time of an activity, and levels of flexibility.

Table 3
Categories and flexibility profiles for activities in the MTMC.

Activity	Category	Flexibility profile ^a	
		Start	Duration
Work	Mandatory ^b	Early: NF	Short: NF
Education		Late: MF	Long: NF
Business trip	Maintenance	Early: MF	Short: MF
Errands, use of services		Late: MF	Long: F
Escort	Discretionary	Early: F	Short: F
Home ^c		Late: MF	Long: F
Shopping		Early: F	Short: F
Leisure		Late: MF	Long: F

^a F = Flexible, MF = Moderately flexible, NF = Not flexible.

^b In this example, we use the term *mandatory* to refer to non-flexible activities with high utilities.

^c Not including mandatory home stays *dawn* and *dusk*.

Table 4
Penalty values by flexibility, in units of utility.

Deviation	Flexibility	Penalty θ
Early start	Flexible (F)	0
	Moderately flexible (MF)	-0.61
	Not flexible (NF)	-2.4
Late start	F	0
	MF	-2.4
	NF	-9.6
Short duration	F	-0.61
	MF	-2.4
	NF	-9.6
Long duration	F	-0.61
	MF	-2.4
	NF	-9.6

activity such as work or education. The leisure activity is included in the all three of the generated schedules, but with varying durations. When included, the escort and errands activities stay relatively close to the preferences.

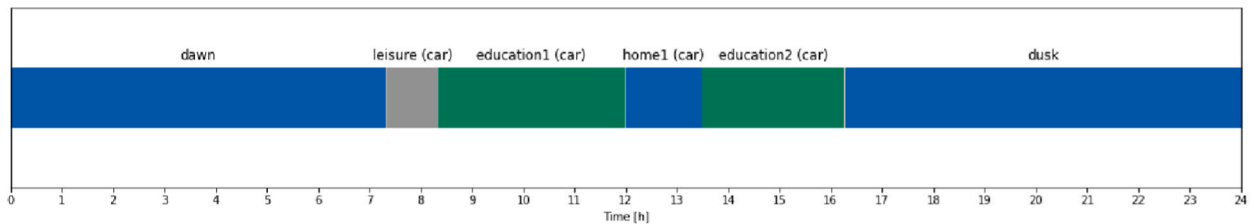
For the choice of transportation mode, none of the solutions include the public transport option. This indicates a consistently higher attractiveness of the car mode for the given parameters.

These results show that for the parameters used in this study the variations in solutions affect mainly the discretionary activities, which have lower penalties for schedule deviations than less flexible activities. Note that we have selected only a small number of unique solutions out of all the generated solutions. The heterogeneity of the solution space (i.e. the distribution from which schedules are drawn) is driven by the relative values of the parameters and the error terms. More specifically, very high penalties (compared to the error variances) lead to semi deterministic problems where the scheduler will consistently output very similar (or the same)

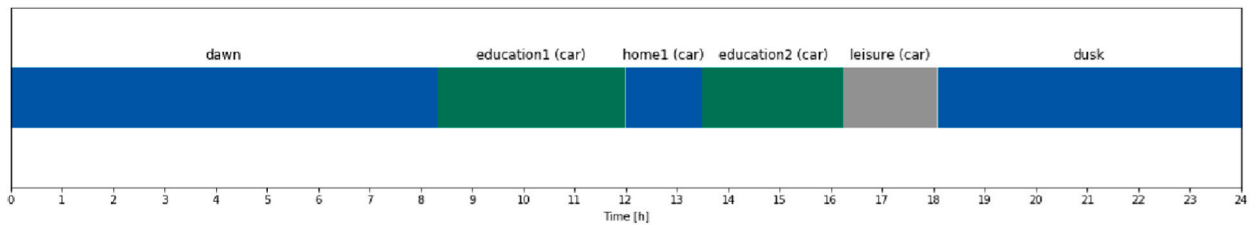
Table 5
 Considered activities and preferences for each individual.

Person	Activity	x_a^e (hh:mm)	τ_a^e (hh:mm)	Location ^a	Mode
Alice	Education (AM)	8:20	3:40	Campus	Car
	Education (PM)	13:30	2:45	Campus	Car
	Education (AM)	8:20	3:40	Campus	PT
	Education (PM)	13:30	2:45	Campus	PT
	Leisure	17:10	0:50	Campus	Car
Bryan	Education	7:30	4:40	Campus	Car
	Shopping	16:30	2:00	Downtown	Car
	Shopping	16:30	2:00	Campus	Car
Claire	Work (A)	14:25	4:25	Office	Car
	Work (B)	14:25	4:25	Office	PT
	Work (C)	14:25	4:25	Library	Car
	Errands	9:45	0:15	Chemist	Car
	Escort	14:10	0:01	Downtown	Car
	Leisure	8:00	1:00	Downtown	Car
	Shopping	13:00	2:00	Shop	Car
Dylan	Escort (Afternoon)	15:10	0:50	School	Car
	Errands	16:40	1:50	Shop	Car
	Escort (Evening)	18:50	0:03	School	Car
	Leisure	19:20	1:30	Gym	Car
	Leisure	19:20	1:30	Gym	Cycling

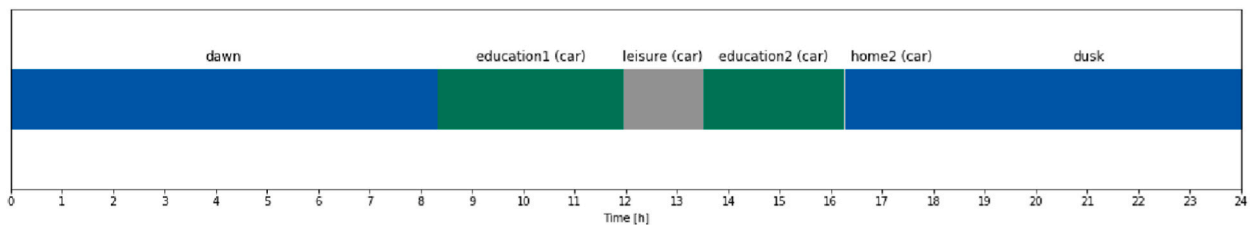
^a Each location is assigned unique coordinates for which travel times are estimated.



(a) Solution 1



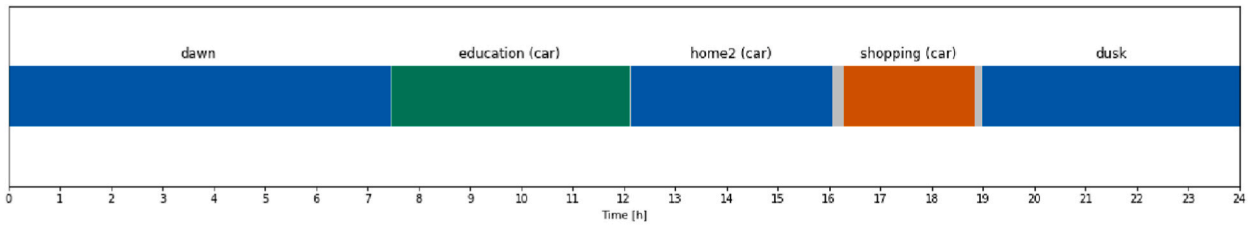
(b) Solution 2



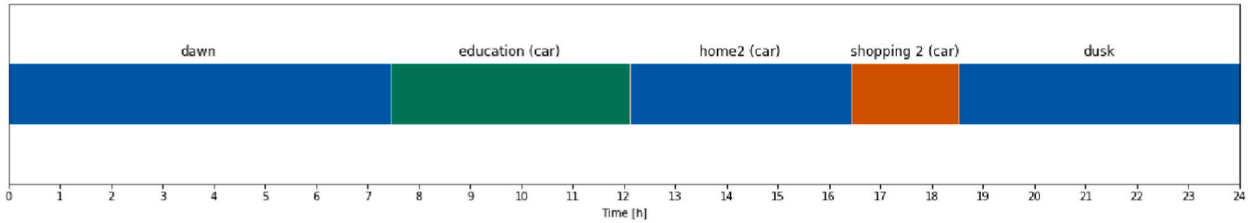
(c) Solution 3

Fig. 4. Generated schedules for Alice.

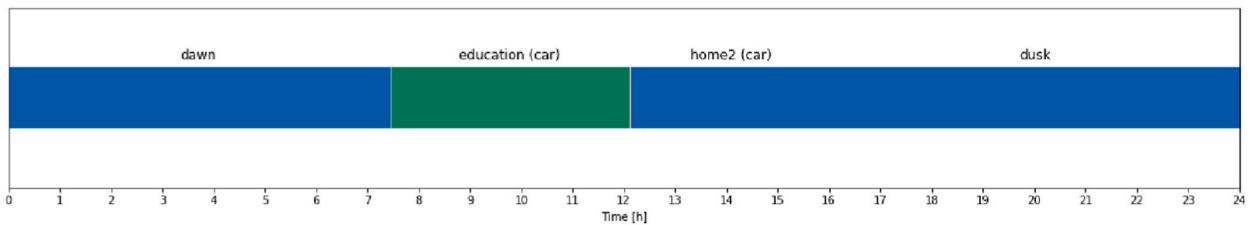
schedules. On the other hand, error terms with very high variance (compared to the penalties) will lead to a diverse set of solutions. An appropriate scale for the error terms must therefore be determined such that the model can generate solutions that are both varied and meaningful.



(a) Solution 1



(b) Solution 2



(c) Solution 3

Fig. 5. Generated schedules for Bryan.

5.4. Distributions of schedules

As mentioned in Section 4, the outcome of the framework is a series of realisations of schedules. To illustrate this concept, we return to the example of Claire, presented in the previous section.

Our framework is designed to capture the interactions of the activities in the schedule, leading to complex distributions of activity participation, start-time, and duration for each considered activity. To show the ability of the framework to draw schedules from a continuous distribution, we repeatedly draw different values of the error terms, and generate the optimal associated schedule. The distribution is then compared to the empirical distribution in the data. We illustrate this example using Claire, and make 1 000 different draws from the schedule distribution. The results are shown in Figs. 8–11.

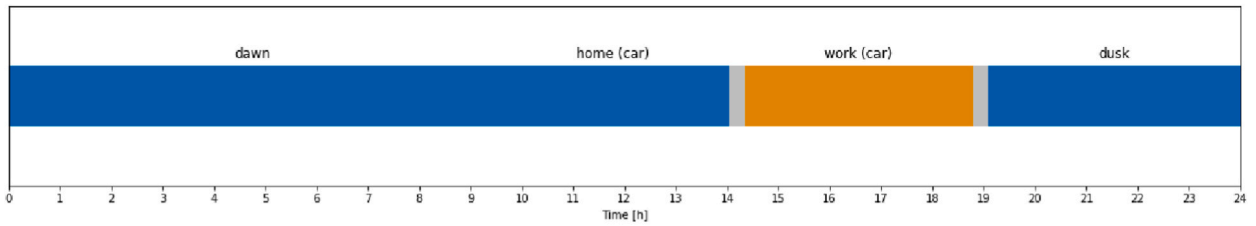
Regarding activity participation (Fig. 8), two types of schedules can be defined: in-home and out-of-home schedules. The former are schedules in which no out-of-home activity (i.e. activity requiring a journey to its destination) is scheduled. Similarly, out-of-home schedules contain at least one trip. 37.1% of the generated schedules contain out-of-home activities (as opposed to a full day spent at home). In the out-of-home schedules, *work* is among the most scheduled activities. Out of the discretionary activities, only *escort* is present in about 20% of the out-of-home schedules. The other activities are almost never scheduled, likely due to their high participation cost.

The simulation results are driven by the random quantities in the utility function, such as the error term ϵ_S .

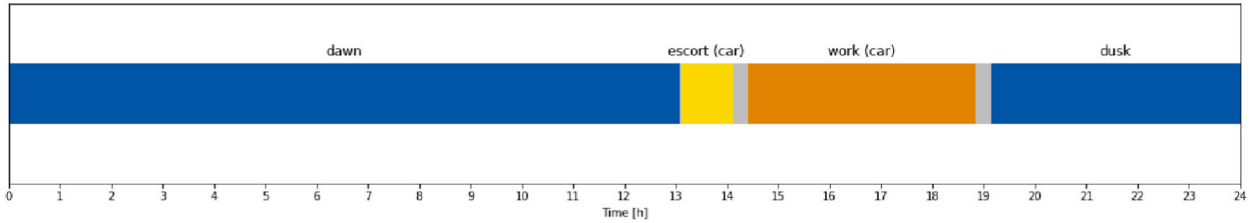
To simulate the effect of the variance of the error term on the activity participation, we fix the variances of the error terms $\sigma^2 = \{1, 10\}$ and generate 1 000 schedules for each value. The ratio of schedules containing each activity is then computed, and shown in Fig. 9. Schedules generated with a higher variance include more frequently discretionary activities. Indeed, we expect the variance to have a large enough magnitude to mitigate the exclusionary effect of the participation costs.

The distributions of start times (Fig. 10) for each activity over 1 000 runs show different profiles, but all seemingly biased towards the desired start time — with the exception of the “home” activity.⁴ The distribution of *work* appears unimodal, centred around the desired start time with very low variance. This is due to the high penalties associated with schedule deviations for this activity (cf. $\mathcal{S}4$).

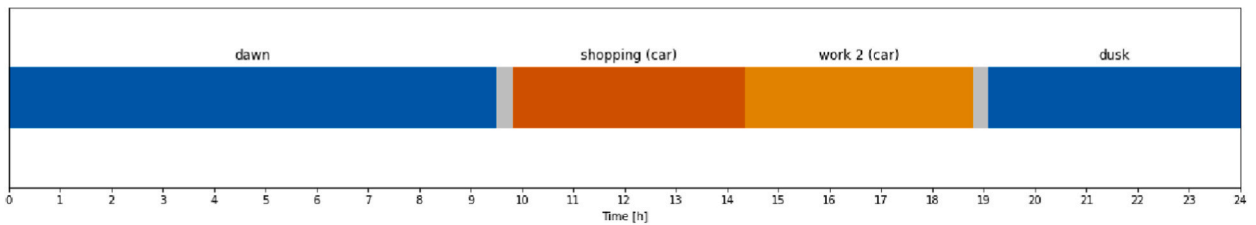
⁴ Regarding the *home* activity, given the constraint that the day must start and end at home, we only show the time of the last return home.



(a) Solution 1



(b) Solution 2



(c) Solution 3

Fig. 6. Generated schedules for Claire.

It is worth noting that the *escort* and *shopping* activities are close to but not centred around the desired time. Given that these are the two activities that had conflicting timings, this result shows the trade-off made during the optimisation process: in most schedules, these activities are started earlier to accommodate other activities for which the penalties for schedule deviations are higher.

Similar observations can be made for the distributions of durations (Fig. 11): the duration assigned to *work* is almost deterministic, with very low variance, and centred around the desired duration, while the durations allocated to flexible activities are more dispersed. Again, when the desired durations involve schedule conflicts, the distributions are not centred around the desired duration, and tend to have large spread (e.g. *escort*).

These distributions are also affected by the random terms. For instance, Fig. 10 shows that increasing the variance of the error terms to 10 does not significantly impact the distribution of start times for the *work* activity. On the other hand, the *escort* activity is more spread in time. As previously noted, the *leisure* and *shopping* activities are more often scheduled, leading to greater variety in the generated solutions. However, all distributions still seem to have a mode relatively close to, or centred around, the desired start time.

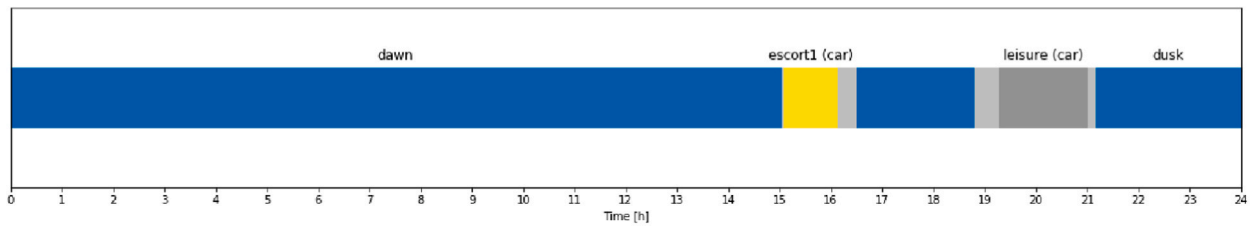
The experimental results show that the framework is able to generate different realisations of chosen schedules for given sets of considered activities, locations, modes, and timing preferences.

The multimodal distributions of the decision variables (start times and durations) highlight the scheduling trade-offs that are made during the optimisation process. These variations impact “flexible” activities in particular, which are characterised by lower penalties for schedule deviations.

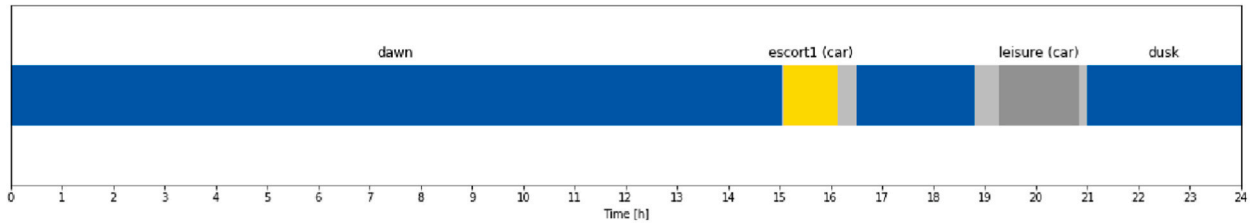
Furthermore, the distributions emphasise the influence of the parameters of the model on its outputs, and consequently, the importance of selecting ranges of values that ensure both varied and stable solutions.

We now compare a distribution generated by our model and the true empirical distribution. We have selected a small subset of individuals from the Microcensus (235 students living in Lausanne) and plotted the schedule frequency (given by the frequency of activity scheduled at a given time of day). We have collected the activities and locations from this sample. The modes of the distributions of start times and duration for each activity were computed and used as proxies for the desired start times and durations input required by the optimisation model. The *education* activity has been set to a lower flexibility level than the other activities ($k = \text{Not Flexible}$). The values of each parameters are summarised in Table 4. We have run several iterations of the model and aggregated the results to obtain a schedule distribution over a day. The empirical and simulated distributions are shown in Fig. 12.

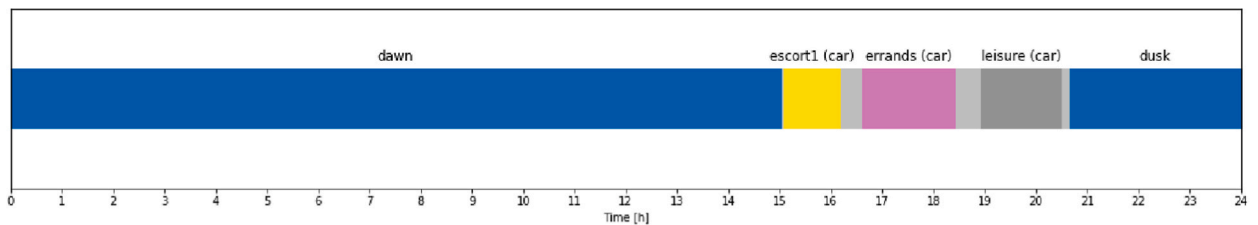
In the empirical distribution, education takes up most of the day, from 7:00 to 22:00, with a higher frequency between 8:00–12:00



(a) Solution 1



(b) Solution 2



(c) Solution 3

Fig. 7. Generated schedules for Dylan.

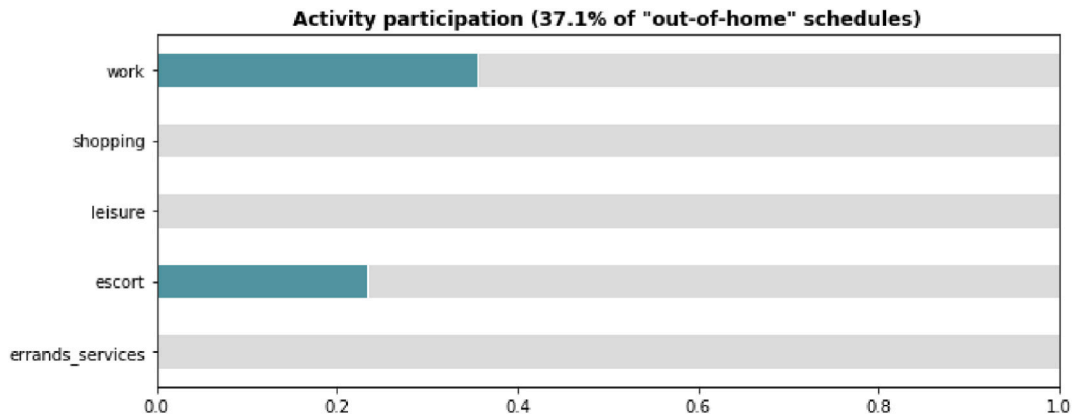


Fig. 8. Proportion of activity participation in out-of-home schedules (1 000 runs).

and 13:00–16:00. The second most frequent activity is *leisure*, which is spread throughout the day between 9:00–4:00 (frequency between 20 and 30% for each hour between 12:00 and 21:00). A small proportion of students work or go shopping during the day. In the simulated distribution, there are in general less out-of-home schedules (i.e. schedules that contain at least one out-of-home activity), but the *education* activity presents a similar profile to the empirical one. The flexible activities such as *leisure* or *shopping* have not been simulated as well: *leisure* is not scheduled after 16:00, and *shopping* is present in greater proportions in the simulated schedules than in the observed ones.

This result shows that the simulator is able to generate a reasonable distribution of schedules for a given population. The parameters describing the problem (flexibility, penalties, desired timings, ...) need to be fine-tuned in order to properly capture the trade-

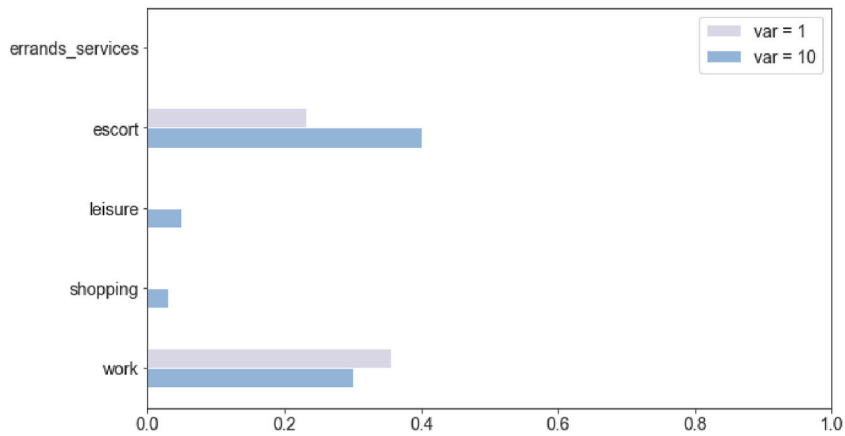


Fig. 9. Distribution of activity participation, for different variances of the random term.

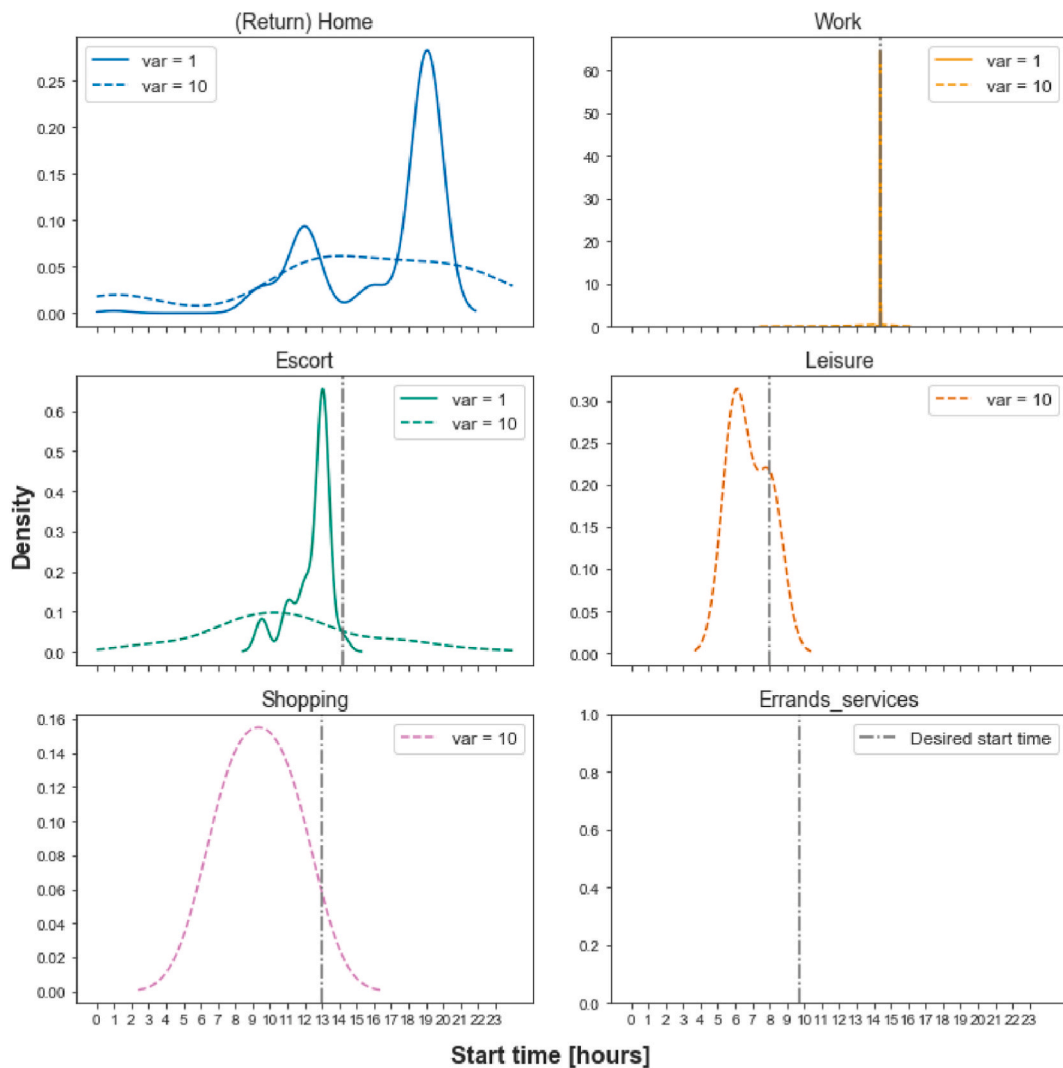


Fig. 10. Distribution of start times per activity and variance of the random term.

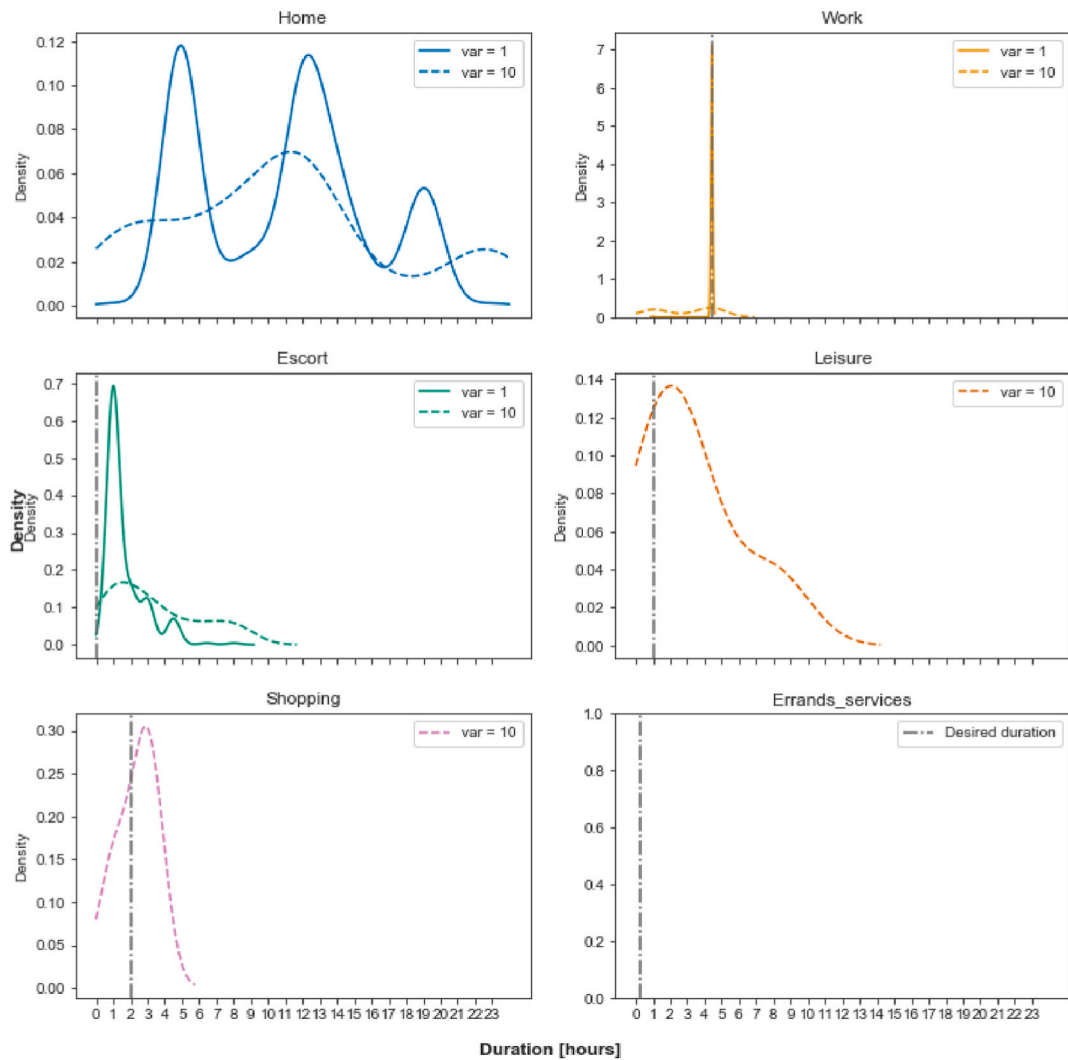


Fig. 11. Distribution of duration per activity and variance of the random term.

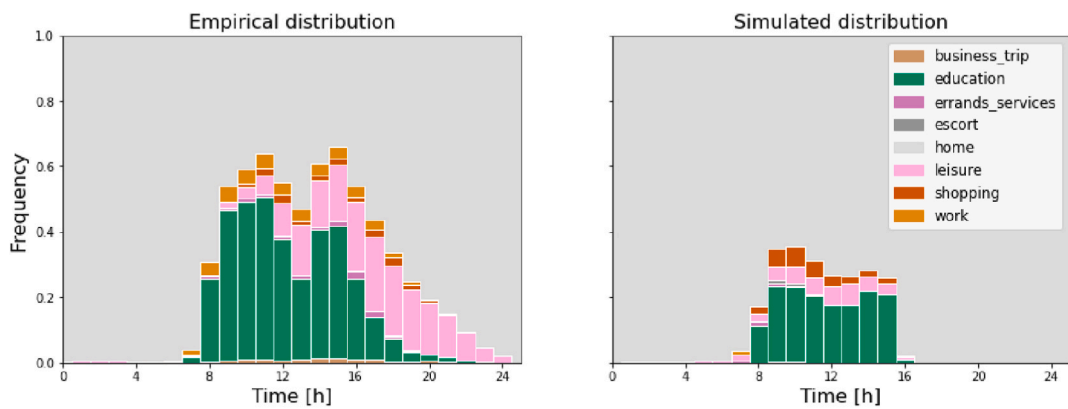


Fig. 12. Comparison of empirical and simulated distribution of activities, for Lausanne students (235 individuals).

offs between activity, especially regarding flexible activities.

6. Conclusion and future work

This paper presents an integrated framework to model the trade-offs made by individuals when scheduling activities. The main characteristics of our methodology are as follows:

- All choices pertaining to daily mobility (activity scheduling, mode choice, activity location) can be considered simultaneously, and trade-offs between these choices are easily modelled.
- A schedule is associated with a utility, consistently with random utility theory (Manski, 1977).
- The choice of a schedule is explicitly modelled as a mixed integer optimisation problem solved by the decision maker.
- Due to the complexity of the choice model, there is no close form probability formulation. Instead, the framework allows the empirical distribution of the choice model to be estimated using simulation.

In its current state, the framework presents several strengths and weaknesses. As a first application of the framework, in this paper we have used a specification of the utility which includes only activity- and travel-specific variables. A linear impact on the utility has been assumed for each of them. Such a simple formulation may not be enough to capture complex behaviours and interactions.

One strength of the framework is its flexibility, which allows the modeller to increase the complexity of the representation without decreasing the practicality of the formulation. As demonstrated with the inclusion of constraints on mode choice (Section 4), many extensions of the model can be implemented in a straightforward way by adding or removing constraints, or by modifying the objective function (utility). While this might increase the computational expense, it does not require any technical or methodological change to accommodate the additions, and the results can still be interpreted from first principles. This characteristic makes the framework particularly interesting for practical applications, or to be integrated into a larger modelling environment with predefined inputs and constraints.

The representation of the modelling elements has also been simplified. Using single combinations of activity, mode of transport and destination as the unit of analysis offers the advantage of predicting different choices of each of the dimensions simultaneously. However, the generation of these combinations relies heavily on assumptions that can limit the ability of the model to deal with specific cases. For this reason, the current implementation does not allow for strong dependencies between activities (e.g. bundles of activities). A careful consideration of how activities are represented in the context of our framework is necessary, with the limitation of available data to confirm our hypotheses.

The search for an optimal exact solution also comes with its set of limitations: the performance and speed of the model depend on its complexity, and the estimation times can quickly become prohibitive for an implementation in practice, that often deal with vast synthetic populations and large amounts of data. Heuristics can be used to reach a solution in a shorter amount of time - but the issue of the validity of the resulting schedule remains. In further work, alternative formulations of the optimisation framework relying on constraint programming may be explored. This work will specifically focus on maintaining the flexibility of the current approach.

In the above case study, we have assumed that the parameters of the model were known. It is usually not the case in practice: they must be estimated from data, using, for instance, maximum likelihood estimation. One significant challenge for the application of maximum likelihood estimation to the activity-based context is the combinatorial nature of the choice set. As the alternatives, or possible schedules, cannot be enumerated, it is necessary to rely on samples of alternatives to estimate the model. This method is well documented in the literature (McFadden, 1978; Guevara and Ben-Akiva, 2013).

Furthermore, the coordination of the activity scheduling decisions among all members of the same household is currently ignored by our framework. An extension that accounts for intra-household interactions is also an interesting topic for future research.

The framework can also be extended to accommodate other complex interactions such as habit formation and its influence on scheduling decisions. Including activity frequency and time of last activity participation as attributes (as done, for example, by Adamowicz (1994), or Arentze and Timmermans (2009)) in the schedule utility function could be a promising avenue.

Finally, the combination of the activity scheduling decisions of many travellers in an area has an impact on the travel patterns of this area. As such, travel times are not exogenous, in contrast to what is assumed by our framework. To address this, the framework could be coupled with mobility simulation tools, that take activity schedules as input, and generate indicators such as travel time and level of congestion as output (e.g. MATSIM (Axhausen et al., 2016), SUMO (Lopez et al., 2018), SimMobility (Adnan et al., 2016), etc.), as shown in Fig. 1. The framework has been successfully integrated within a long-term transport demand framework developed by the Swiss Federal Railways (Manser et al., 2021; Scherr et al., 2019).

CRediT authorship contribution statement

Janody Pougala: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Tim Hillel:** Conceptualization, Methodology, Investigation, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Michel Bierlaire:** Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration.

Declaration of competing interest

We have no conflict of interest to disclose.

Acknowledgments

This work is part of a collaborative project with the Swiss Federal Railways (SBB), supported by the Swiss Innovation Agency (Innosuisse) (Project number: 2155006680).

Michel Bierlaire dedicates this article to John Polak, who introduced him to the fascinating world of travel behaviour in the early days of his PhD.

References

- Adamowicz, W., 1994. habit formation and variety seeking in a discrete choice model of recreation demand. *J. Agric. Resour. Econ.* 19.
- Adler, T., Ben-Akiva, M., 1979. A theoretical and empirical model of trip chaining behavior. *Transp. Res. Part B Methodol.* 13 (3), 243–257.
- Adnan, M., Pereira, F.C., Azevedo, C.M.L., Basak, K., Lovric, M., Raveau, S., Zhu, Y., Ferreira, J., Zegras, C., Ben-Akiva, M., 2016. SimMobility: a multi-scale integrated agent-based simulation platform. In: 95th Annual Meeting of the Transportation Research Board Forthcoming in Transportation Research Record.
- Arentze, T., Timmermans, H., 2000. Albatross: a learning based transportation oriented simulation system. Technische Universiteit Eindhoven / EIRASS.
- Arentze, T.A., Timmermans, H.J., 2009. A Need-Based Model of Multi-Day, Multi-Person Activity Generation. *Transportation Research Part B, Methodological.*
- Arentze, T.A., Ettema, D., Timmermans, H.J., 2011. Estimating a model of dynamic activity generation based on one-day observations: method and results. *Transp. Res. Part B Methodol.* 45 (2), 447–460.
- Axhausen, K., Horni, A., Nagel, K., 2016. *The Multi-Agent Transport Simulation MATSim*. Ubiquity Press.
- BFS, 2007. *Die Einkommens- und Verbrauchserhebung mit neuer Methode*. Number 343561. Bundesamt für Statistik (BFS), Neuchâtel.
- BFS, 2021. *Kosten und Finanzierung des Verkehrs 2018*. Number 18224159. Bundesamt für Statistik (BFS), Neuchâtel.
- Bhat, C.R., 2005. A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions. *Transp. Res. Part B Methodol.* 39 (8), 679–707.
- Bhat, C.R., 2018. A new flexible multiple discrete–continuous extreme value (MDCEV) choice model. *Transp. Res. Part B Methodol.* 110, 261–279.
- Bhat, C.R., Guo, J.Y., Srinivasan, S., Sivakumar, A., 2004. Comprehensive econometric microsimulation for daily activity-travel patterns. *Transport. Res. Rec.: J. Transport. Res. Board* 1894 (1), 57–66.
- Bowman, J.L., Ben-Akiva, M.E., 2001. Activity-based disaggregate travel demand model system with activity schedules. *Transport. Res. Pol. Pract.* 35 (1), 1–28.
- Bradley, M., Bowman, J., 2008. Design features of activity-based microsimulation models for U.S. Metropolitan planning organizations: a summary. In: *Transportation Research Board Conference Proceedings*, pp. 11–20.
- Castiglione, J., Bradley, M., Gliebe, J., 2014. *Activity-Based Travel Demand Models: A Primer*. Transportation Research Board, Washington, D.C.
- Chapin, F.S., 1974. *Human Activity Patterns in the City: Things People Do in Time and in Space*, tome 13. Wiley-Interscience.
- Charypar, D., Nagel, K., 2005. Generating complete all-day activity plans with genetic algorithms. *Transportation* 32 (4), 369–397.
- Ettema, D., Borgers, A., Timmermans, H., 2000. A simulation model of activity scheduling heuristics: an empirical test. *Geogr. Environ. Model.* 4 (2), 175–187.
- Ettema, D., Bastin, F., Polak, J., Ashiru, O., 2007. Modelling the joint choice of activity timing and duration. *Transport. Res. Pol. Pract.* 41 (9), 827–841.
- Golledge, R.G., Kwan, M., Garling, T., 1994. Computational-process modelling of household travel decisions using a geographical information system. UC Berkeley: University of California Transportation Center.
- Guevara, C., Ben-Akiva, M., 2013. Sampling of alternatives in logit mixture models. *Transp. Res. Part B Methodol.* 58, 185–198.
- Hägerstrand, T., 1970. What about people in regional science? *Pap. Reg. Sci.* 24 (1), 6–21.
- Hilgert, T., Heilig, M., Kagerbauer, M., Vortisch, P., 2017. Modeling week activity schedules for travel demand models. *Transport. Res. Rec.: J. Transport. Res. Board* 2666 (1), 69–77.
- Lopez, P.A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.-P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P., Wießner, E., 2018. Microscopic traffic simulation using sumo. In: *The 21st IEEE International Conference on Intelligent Transportation Systems*. IEEE, pp. 2575–2582.
- Manser, P., Haering, T., Hillel, T., Pougala, J., Krueger, R., Bierlaire, M., 2021. Resolving Temporal Scheduling Conflicts in Activity-Based Modelling. Technical Report TRANSP-OR 211209, Transport and Mobility Laboratory. Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.
- Manski, C.F., 1977. The structure of random utility models. *Theor. Decis.* 8 (3), 229–254.
- McFadden, D., 1978. *Modeling the Choice of Residential Location*, vol. 673. Transportation Research Record.
- Miller, E., Roorda, M., 2003. Prototype model of household activity-travel scheduling. *Transport. Res. Rec.* 1831 (1), 114–121.
- Nurul Habib, K.M., 2011. A random utility maximization (RUM) based dynamic activity scheduling model: application in weekend activity scheduling. *Transportation* 38 (1), 123–151.
- Nurul Habib, K., 2018. A comprehensive utility-based system of activity-travel scheduling options modelling (CUSTOM) for worker's daily activity scheduling processes. *Transportmetrica: Transport. Sci.* 14 (4), 292–315.
- Nurul Habib, K.M., Miller, E., 2009. Modelling Activity Generation: A Utility-Based Model for Activity-Agenda Formation. *Transportmetrica*.
- Office fédéral de la statistique and Office fédéral du développement Territorial, 2017. *Comportement de la population en matière de transports. Résultats du microrecensement mobilité et transports 2015*. Technical report. Neuchâtel, Berne.
- Palma, D., Enam, A., Hess, S., Calastri, C., Crastes dit Sourd, R., 2021. Modelling multiple occurrences of activities during a day: an extension of the MDCEV model, 9 (1), 456–478. <https://doi.org/10.1080/21680566.2021.1900755>.
- Pas, E., 1985. State of the art and research opportunities in travel demand: another perspective. *Transport. Res. Gen.* 19 (5–6), 460–464.
- Pendyala, R., Ryuichi, K., Akira, K., Toshiyuki, Y., Satoshi, F., 2005. *Florida Activity Mobility Simulator: Overview and Preliminary Validation Results*, vol. 1921. Transportation Research Record, pp. 123–130.
- Recker, W.W., Root, G.S., 1981. *Toward a Dynamic Model of Individual Activity Pattern Formulation*.
- Recker, W., McNally, M., Root, G., 1986. A model of complex travel behavior: Part I—theoretical development. *Transport. Res. Gen.* 20 (4), 307–318.
- Scherr, W., Joshi, C., Manser, P., Frischknecht, N., Métrailler, D., 2019. An activity-based travel demand model of Switzerland based on choices and constraints. In: *hEART 2019: 8th Symposium of the European Association for Research in Transportation*.
- Schmid, B., Molloy, J., Peer, S., Jokubauskaite, S., Aschauer, F., Hössinger, R., Gerike, R., Jara-Diaz, S.R., Axhausen, K.W., 2021. The value of travel time savings and the value of leisure in Zurich: estimation, decomposition and policy implications. *Transport. Res. Pol. Pract.* 150, 186–215.
- Small, K.A., 1982. *The Scheduling of Consumer Activities: Work Trips*. American Economic Review.
- Smith, L., Beckman, R., Baggerly, K., 1995. *TRANSIMS: Transportation Analysis and Simulation System*. Technical Report. Los Alamos National Lab., NM (United States).
- Timmermans, H., 2003. Recent and expected directions in modelling transportation. *SATC* 2003.
- Vovsha, P., Bradley, M., Bowman, J., 2005. Activity-based travel forecasting models in the United States: progress since 1995 and prospects for the future. In: *EIRASS Conference on Progress in Activity-Based Analysis*.
- Weis, C., Kowald, M., Danalet, A., Schmid, B., Vrtic, M., Axhausen, K.W., Mathys, N., 2021. Surveying and analysing mode and route choices in Switzerland 2010–2015. *Travel Behav. Soc.* 22, 10–21.