From one-day to multiday activity scheduling: Extending the OASIS Framework

Janody Pougala

Tim Hillel

Michel Bierlaire

STRC conference paper 2023 **10-12th May, 2023**

STRC | **23rd Swiss Transport Research Conference**
Monte Verità / Ascona, May 10-12, 2023

From one-day to multiday activity scheduling: Extending the OASIS Framework

Janody Pougala
TRANSP-OR
EPF Lausanne
janody.pougala@epfl.ch

Tim Hillel
Behaviour and Infrastructure Group
University College London
tim.hillel@ucl.ac.uk

Michel Bierlaire
TRANSP-OR
EPF Lausanne
michel.bierlaire@epfl.ch

10-12th May, 2023

Abstract

Applications of activity-based models for the estimation of transport demand have demonstrated to achieve greater behavioural realism than traditional trip-based models. However, state-of-the art models focus on single-day schedules as focal points to their estimations, thus ignoring fundamental dynamics that explain individual behaviour over longer periods of time. Several authors have highlighted the importance of multiday analyses in activity-travel contexts, which are still lacking in many state-of-the-art framework. In this paper, we present an extension of the OASIS framework, an integrated model for the simulation of single day schedules, to include intrapersonal interactions influencing longer term decisions. We formulate the multiday problem as a multiobjective optimisation problem where each day d is associated with a utility U_d. We consider an activity-based set up where individuals maximise the total utility of their schedules over multiple days (e.g. week). We discuss implications and requirements of this formulation, and illustrate the methodology with chosen examples.

Keywords

activity-based modelling, scheduling, multiday, transport demand

Suggested Citation

This work was supported by an EPFL ENAC Interdisciplinary Cluster Grant. The authors would like to thank Luca Bataillard for the Constraint Programming implementation of OASIS, and the IVT-ETHZ group for providing the MOBIS dataset.

Contents

List of Tables

List of Figures

1 Introduction

Activity-based models (ABMs) stem from the fundamental assumption that travel demand is derived from the need to perform activities, rooted in a spatiotemporal context, and influenced by personal and environmental factors. By focusing on individuals and explicitly considering these interactions, ABMs aim to be more behaviourally realistic than traditional trip-based models, and to provide more flexible and targeted insights on individual mobility. Successful applications of activity-based models have demonstrated the added value of shifting the focus to individuals and their activities. Two main approaches can be cited: utility-based and rule-based models. The scheduling process is a result of random utility maximisation for the former, and the satisfaction of a set of spatio-temporal rules for the latter. While historically both approaches were considered contradictory, there has been increased research in hybrid models which combine elements from either theory.

In practice, most operational activity-based models have focused on single-day analyses. This common simplifying assumption greatly limits the behavioural realism of the models, as they are unable to properly capture the dynamics and processes involved in the scheduling of activities over multiple days. Decisions taken daily are affected by both *habits* built over time, and *forward-looking behaviour* (Bierlaire *et al.*, 2021), where individuals decide based on the expected outcomes of future decisions. In addition, some constraints are not necessarily applicable on a 24h period, but on longer periods of time (e.g. shopping frequency which becomes more constraining the longer the individual goes without necessities). In order to realistically model the activity-travel behaviour, it is crucial to explicitly integrate these elements.

In this paper, we present and discuss assumptions and methodological requirements to extend a single-day scheduling framework to a multiday scope. We focus particularly on the OASIS framework. OASIS is an integrated framework to simulate daily activity schedules by considering all choice dimensions (activity participation, timing decisions, mode and location choice) simultaneously (Pougala *et al.*, 2022). The current implementation is an optimisation model where the schedule utility of a single day is maximised, subject to constraints (including a 24h time budget). We do not intend in this paper to present an operational multiday activity scehduler, but instead discuss different avenues to be explored to address this issue: including methodological extensions, modelling assumptions, and data requirements.

In Section 2, we present relevant research on the development of multiday activity-based models. In Section 3, we introduce the single-day OASIS framework, the assumptions and methodology to accommodate multiday considerations into the model. Finally, we illustrate the preliminary modelling additions with selected examples from the MOBIS dataset (Molloy *et al.*, 2022).

2 Literature review

Many authors have identified a significant limitation of current activity-based models: the majority of models focus on the simulation of a single day, thus ignoring significant dynamics and correlations that arise from the scheduling over multiple days or longer periods of time (e.g. Roorda and Ruiz, 2008; Calastri *et al.*, 2020). However, it is agreed that there exist specific intrapersonal factors that influence the scheduling process. Roorda and Ruiz (2008) formulate the hypotheses that a person activity/travel planning behaviour depends on their behaviour on other days of the week, and that there are two main components to this dynamic behaviour:

1. same-day and next-day substitution effects for activities and trips, and
2. latent propensity to engage in some activities or to choose a specific mode of transportation.

One reason for the lack of research on multiday dynamics is the lack of available longitudinal data, which contains sufficient information to calibrate a multiday ABM. This leads several authors to propose methodologies to derive intrapersonal indicators from single-day trip data (e.g. Arentze *et al.*, 2011; Hilgert *et al.*, 2017). It is an effective solution to circumvent the data issue, but can introduce significant biases in the outputs as many behavioural assumptions must be taken with limited possibilities of validating them.

In ABM research, most studies on intrapersonal variability focus on specific aspects of the scheduling process but not on the activity-travel behaviour as a whole (Zhang *et al.*, 2021). For example, some authors investigate the impact of multiday dynamics on activity generation only (e.g. Nurul Habib and Miller, 2008; Arentze *et al.*, 2011).

Nurul Habib and Miller (2008) address the issue of activity generation (i.e. generating sets of feasible activities to be scheduled), for a week time-frame. They solve a utility-based model for each day of the week, and include in the utility functions for each day the frequency of participation for each activity on the previous day. Their assumption is that the activities that are considered for each day are influenced by the activities that were actually performed previously. They find that estimating individual daily models with previous-day effect yields better model fit than estimating an aggregated model over the week.

Arentze *et al.* (2011) also introduce multiday considerations for dynamic activity generation. More specifically, they propose a methodology to generate activities for multiple days by using one-day observations from trip diary data. They assume that each observation is a draw from a long-term distribution of activity patterns. They perform random utility maximisation, and consider that the utilities of the activity patterns are dependent on the time elapsed since the last performed activity. In addition, they postulate the existence of individual preferences

On the other hand, Zhang *et al.* (2021) model intrapersonal (day-to-day) variability of full activity-travel patterns, represented as spatio-temporal networks, with a bi-level multinomial logit model (MNL). The first level models the utility of choosing a representative pattern with respect to the day of the week, while the second level explains the alternative-specific constants of the first level with socio-demographic characteristis.

One common limitation of these models is that the multiday dynamics depend solely on the past, and decisions for future days are ignored. In addition, when multiple days are considered, the focus remains on the correlations between days, and scheduling dynamics within a day take less priority.

Calastri *et al.* (2020) argue that both within and between days correlations must be taken into account in order to develop realistic multiday activity-based models. They propose different specifications of the MDCEV (Multiple Discrete-Continuours Extreme Value model) in order to integrate and test the effects of these correlations. In their first specification, they consider day-specific non-additive utility functions and time constraints. In the second case, they consider correlated and additive utility functions.

In this research, we propose some considerations to develop a methodology for the simulation of multiday activity-travel scheduling, within the scope of the OASIS framework. This framework integrates all choice dimensions simultaneously in order to capture scheduling trade-offs, which implies a significant level of complexity when the model is scaled up to simulate multiple correlated days.

3 Methodology

3.1 OASIS framework

In this section, we briefly introduce the OASIS framework (Pougala *et al.*, 2022). The framework (Fig. 1) is composed of two main elements:

1. A simulation model that outputs distributions of feasible schedules for given individuals, based on a mixed-integer optimisation model where the objective function is the utility of the schedule.
2. An estimation component to calibrate the parameters of the said utility function.

The methodology of the simulation framework is built upon first behavioural principles:

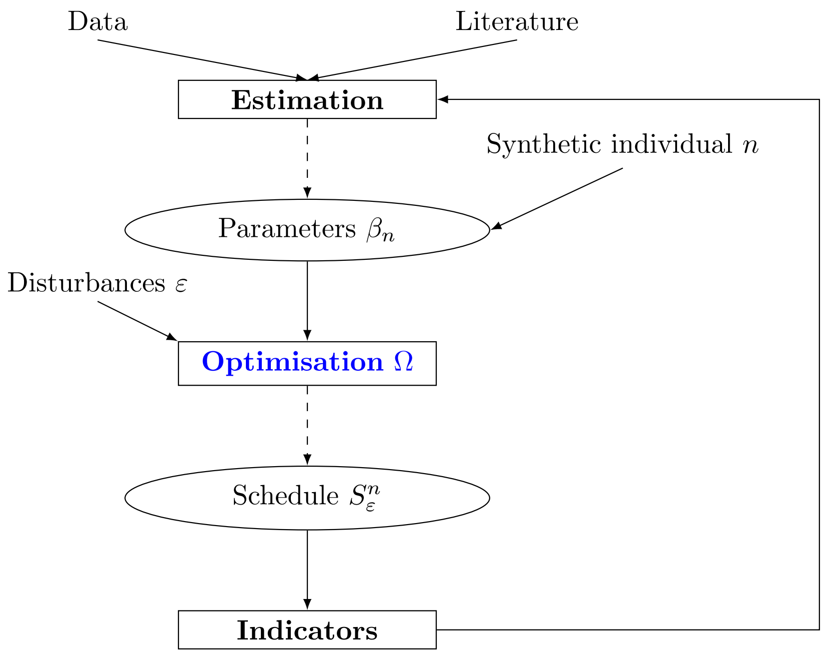
1. Individuals have a time budget which constrains their activity schedules. Considering both in and out of home activities, the total duration of the schedule must be equal to the available budget.
2. Each considered activity a is associated with a utility U_{an}. This utility translates the satisfaction derived by the individual by performing the activity. We assume this utility to be time-dependent.
3. The scheduling process itself is assumed to be driven by the desire to maximise the total satisfaction, or utility, provided by the activities subject to the given time budget constraint.

3.1.1 Simulation

The scheduling process for a given individual n, a set of activities to be scheduled A_n, and sets of possible modes M_n and $L_{a,n}$ is summarised in 1. The set of estimated parameters $\beta_{a,n}$ is also provided as input.

The objective function of the maximisation problem is the utility function of the schedule. This function is expressed as a sum of a deterministic element and an error term (Eq.1).

Figure 1: OASIS framework



We assume the distribution of the error term to be known.

$$U_S = V_S + \varepsilon_S \tag{1}$$

The decision variables (activity participation, start time, duration, sequence,...) are chosen such as to maximise the utility of the schedule subject to constrained. For a draw of the error term ε^r_S, the maximisation problem becomes deterministic, and yields one optimal schedule S^r which is a draw from the distribution of schedules for n.

The constraints of the optimisation problem can be classified in two categories:

1. Mathematical constraints: they ensure that the resulting schedule is valid. For example, the time budget T cannot be exceeded, activities cannot overlap,...
2. Context-dependent constraints: they are additional rules that influence the activity-travel behaviour. For example, some activities must take precedence over others (e.g. picking up children from school must happen after they were picked up),...

As defined in Pougala *et al.* (2022), the schedule utility U_S is the sum of a generic utility U

Algorithm 1 Simulation of activity schedules

Initialise n, β_n, A_n, M_n, L_n
for r = 1,2,...,R **do**
    Draw ε_S^r from distribution of error terms.
    Draw schedule S_n^r by solving $\Omega = \max U_S(X_n, \beta_n, \varepsilon_S^r)$ s.t. constraints

associated with the whole schedule and utility components capturing the activity-travel behaviour:

$$U_S = U + \sum_{a=0}^{A-1} (U_a^{\text{participation}} + U_a^{\text{start time}} + U_a^{\text{duration}} + \sum_{b=0}^{A-1} U_{a,b}^{\text{travel}}). \tag{2}$$

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 (e.g. resource availability at the level of the household).
2. The utility $U_a^{\text{participation}}$ associated with the participation of the activity a, irrespective of its starting time and duration.

   $$U_a^{\text{participation}} = \gamma_a + \beta_{\text{cost}} c_a + \varepsilon_{\text{participation}}, \tag{3}$$

   where γ_a and β_{cost} are unknown parameters to be estimated from data, and $\varepsilon_p articipation$ is an error term.
3. The utility $U_a^{\text{start time}}$, which captures the perceived penalty created by deviations from the preferred starting time.

   $$U_a^{\text{start time}} = \theta_a^{\text{early}} \max(0, x_a^* - x_a) + \theta_a^{\text{late}} \max(0, x_a - x_a^*) + \varepsilon_{\text{start time}}, \tag{4}$$

   where $\theta_a^{\text{early}} \leq 0$ and $\theta_a^{\text{late}} \leq 0$ are unknown parameters to be estimated from data, and $\varepsilon_{\text{start time}}$ is an error term.
   The first (resp. second) term captures the disutility of starting the activity earlier (resp. later) than the preferred starting time.
4. The utility U_a^{duration} associated with duration. This term captures the perceived penalty created by deviations from the preferred duration.

$$U_a^{\text{duration}} = \theta_a^{\text{short}} \max(0, \tau_a^* - \tau_a) + \theta_a^{\text{long}} \max(0, \tau_a - \tau_a^*) + \varepsilon_{\text{duration}} \tag{5}$$

where $\theta_a^{\text{short}} \leq 0$ and $\theta_a^{\text{long}} \leq 0$ are unknown parameters to be estimated from data, and $\varepsilon_{\text{duration}}$ is an error term. 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 $U_{\text{travel}}^{a,b}$ associated with the trip from ℓ_a to ℓ_b. irrespective of the travel time. This term is composed of the penalty associated with the travel time ρ_{ab}, and other travel variables (including variables such as cost, level of service, etc.) Here, we illustrate the framework with a specification involving travel cost. It also includes an error term, capturing the unobserved variables.

$$U_{a,b}^{\text{travel}} = \beta_{t,\text{time}} \rho_{ab} + \beta_{t,\text{cost}} c_t + \varepsilon_{\text{travel}} \tag{6}$$

where $\beta_{t,\text{time}}$ and $\beta_{t,\text{cost}}$ are unknown parameters to be estimated from data, and $\varepsilon_{\text{travel}}$ is an error term.

The schedules generated by the simulator must be *feasible*, according to a set of constraints defined at the level of the individual or the household by the modeller. For example, a schedule is feasible if:

- it does not exceed the maximum (time or cost) budget,
- each activity starts when the trip following the previous activity is finished,
- trips using mode m are only made if and when m is available,
- each activity meets its respective requirements (e.g. participation of other members of the household, feasible time windows, follows/precedes another activity)

Individuals are assumed to choose the schedule with the highest utility for a given set of parameters, and subject to constraints. The decision variables of the maximisation problem are the following:

- ω_a: binary participation variable,
- z_{ab}: binary succession variable,
- x_a: start time,

- τ_a: duration.

The objective function is derived from (2):

$$\max_{\omega,z,x,\tau} U(\omega, z, x, \tau) + \sum_{a=0}^{A} \omega_a (U_a^{\text{participation}} + U_a^{\text{start time}} + U_a^{\text{duration}}) + \sum_{a=0}^{A} \sum_{b=0}^{A} z_{ab}(U_{a,b}^{\text{travel}}). \quad (7)$$

The constraints are

$$\sum_a \sum_b (\omega_a \tau_a + z_{ab}\rho_{ab}) = T, \quad (8)$$

$$\sum_a \sum_b (\omega_a c_a + z_{ab}\kappa_{ab}) \leq B, \quad (9)$$

$$\begin{aligned}
\omega_{\text{dawn}} = \omega_{\text{dusk}} &= 1, && && (10)\\
\tau_a &\geq \omega_a \tau_a^{\min}, && \forall a \in A, && (11)\\
\tau_a &\leq \omega_a T, && \forall a \in A, && (12)\\
z_{ab} + z_{ba} &\leq 1, && \forall a, b \in A, a \neq b, && (13)\\
z_{a,\text{dawn}} = z_{\text{dusk},a} &= 0, && \forall a \in A, && (14)\\
\sum_a z_{ab} &= \omega_b, && \forall b \in A, b \neq \text{dawn}, && (15)\\
\sum_b z_{ab} &= \omega_a, && \forall a \in A, a \neq \text{dusk}, && (16)\\
(z_{ab} - 1)T &\leq x_a + \tau_a + z_{ab}\rho_{ab} - x_b, && \forall a, b \in A, a \neq b, && (17)\\
(1 - z_{ab})T &\geq x_a + \tau_a + z_{ab}\rho_{ab} - x_b, && \forall a, b \in A, a \neq b, && (18)\\
\sum_{a \in G_k} \omega_a, &\leq 1 && k = 1, \ldots, K, && (19)\\
\alpha_a^m &= 1 && \forall a \in G_{\text{home}} && (20)\\
\omega_a &\leq \alpha_a^m && \forall a \in A^m && (21)\\
\omega_a &\geq \omega_b + z_{ab} - 1 && \forall a \in A, b \in A \setminus G_{\text{home}} && (22)\\
\omega_b &\geq \omega_a + z_{ab} - 1 && \forall a \in A, b \in A \setminus G_{\text{home}} && (23)\\
x_a &\geq \gamma_a^-, && \forall a \in A, && (24)\\
x_a + \tau_a &\leq \gamma_a^+, && \forall a \in A. && (25)
\end{aligned}$$

Equation (8) 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 (9) 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 (10) ensures that each schedule begins and ends with the dummy activities *dawn* and *dusk*. Equations (11) and (12) enforce consistency with the activity duration by requiring the activity to have a duration greater or equal than the minimal duration (**??**) and for the activity to have zero duration if it does not take place. Equations (13)-(17) constrain the sequence of the activities: (13) 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 (14)-(16) state that each activity has only one predecessor (excluding the first activity), and each activity only one successor (excluding the last activity). Equation (17) enforces time consistency between two consecutive activities (with travel time ρ_{ab}). Equation (19) ensures that only one activity within a group of duplicates G is selected. Equations (20)-(23) define the constraints related to the choice of mode of transportation. (20) ensures that all private modes m are always available for activities (or trips) starting from home. (21) only allows alternatives associated with a private mode m to take place if m is available, while (22) and (23) enforce mode consistency between two consecutive activities, excluding returns home where a different (private) mode can be chosen. Finally, (24) and (25) are time-window constraints.

3.2 Multiday analysis

3.2.1 Hypotheses

The first behavioural principles at the core of the single-day framework still hold in the multiday case. The extension to multiday is done by relaxing the assumption that days are scheduled independently, but rather, that each day is planned by considering constraints, preferences and decisions taken within a larger time horizon. There are three main mechanisms which influence decisions over time (Bierlaire *et al.*, 2021):

1. changes in external conditions over time,
2. habitual behaviour,
3. forward-looking planning.

The second and third points are behavioural processes that are specific to each individual.

Habits translate the ability of the decision-maker to learn from past experiences, and influence the preferences and perception of current options. On the other hand, *forward-looking planning* affects current decisions by anticipating future outcomes and their associated utility. For the scope of this research, we focus on the integration of habits to introduce correlation between days. Habitual behaviour and learning can considered by including dedicated terms in the utility function, but also by calibrating the parameters such that the generated schedules match habitual patterns of activities, or activity motifs (Schultheiss, 2021).

3.2.2 New definitions

The multiday extension can be considered on multiple levels:

1. Input:
   - **Preferences**: The assumption that individuals have desired start times x_a^* and τ_a^* for each activity still holds in the multiday case. However, this preference might depend on the day d. For example, an individual may prefer to start working at 08:30 on Mondays, but at 09:00 on Tuesdays because they have to escort their children to school. In addition, we consider the preference for activity frequency f_a^*, which is the number of times that an individual prefers to perform the activity in the time horizon T. We can therefore differentiate regular from occasional activities (e.g. work vs. leisure). We also consider a preferred day d_a^* for participation to activity a. f_a^* and d_a^* are included in the daily utility functions U_S^d.
2. Model:
   - **Time budget**: considering a time horizon T of $D = 1, 2..., d$ days, the time budget can be increased from 24h to 24d hours. This implies that a specific activity a can potentially be scheduled on any day d, depending on the preferences and flexibility of the individual. This reflects more closely activity planning behaviour, by allowing some activities to be scheduled at a later date if it increases the overall utility.
   - **Objective function:** Each day d of the time horizon is associated with a time-dependent utility function U_S^d (Eq. (2)). However, we now solve a multiobjective optimisation problem, where the objective function is the weighted sum of the daily schedules over the time horizon (Eq. (26))

$$\Omega = \max_{\boldsymbol{\omega},\mathbf{z},\mathbf{x},\boldsymbol{\tau}} \sum_d w_d U_S^d \tag{26}$$

- **Decision variables**: the decision variables the same defined for the single day problem. In the multiday case, the decision variables $\boldsymbol{\omega}, \mathbf{z}, \mathbf{x}, \boldsymbol{\tau}$ are vectors of size T with T the number of days (time horizon). Each element of the vector is the decision for day d.
- **Daily weights**: The multiobjective optimisation defined in Eq. (26) is a weighted sum of utilities. The weights translate a priority between days, i.e. we assume that days are not considered equally in terms of activity utility. For example, if we distinguish week days vs. weekends, we can consider that the former have higher priority in terms of scheduling. This implies that deviations from preferences on these higher priority days would be more penalised than on other days.

Table 1 summarises the main methodological differences between the single-day and the multiday simulator.

Table 1: Methodological differences between single-day and multiday framework

Feature	Single day	Multiday
Time horizon T	1 day	d days
Objective function	Schedule utility U_S	Multi-objective $\sum_d w_d U_S^d \; \forall d \in T$
Decision variables	Specific to activities	Specific to days and activities
Constraints	Specific to activities	Specific to days and activities
Preferences	Start time, duration	Daily start time and duration, frequency over T

3.2.3 Utility function

Considering a day d in the time horizon, the multiday utility function is presented in Equation (27).

$$U_S^d = U + \sum_{a=0}^{A-1} (U_{a,d}^{\text{participation}} + U_{a,d}^{\text{start time}} + U_{a,d}^{\text{duration}} + \sum_{b=0}^{A-1} U_{a,b,d}^{\text{travel}}). \tag{27}$$

In particular, the terms for start time and duration are day-specific, which implies that individuals have different sensitivities towards deviations from their preferred timings based on the day of the week. It may also imply that scheduling preferences are also specific to the day.

The objective function of the optimisation problem is also updated as we now solve a multiobjective optimisation problem (Eq. (26)). The weights ω_d translate the assumption that individuals prioritise certain days when making their multiday scheduling decisions. For instance, some days might be more contraining than others, meaning that the overall scheduling process will tend to satisfy constraints of this specific day more thoroughly than other, more flexible days.

The weights are difficult to define as they cannot be directly observed. However, analysing frequency of activity participation over multiple days can provide insights on trends on preferred days.

3.2.4 Constraints

As we assume that multiple days are correlated, the assumption that schedules should start and end at home can be relaxed. Indeed, it is now more likely to observe schedules with overnight activities (e.g. leisure) which would shift the return home of the next day or to a later start time. The *dawn* and *dusk* variables make less sense in this context, and can be redefined in the following way:

1. We introduce the dummy variables *first* and *last*, which respectively identify the first and last day of the time horizon. These variables are helpful to enforce consistency constraints across days: each day except *last* should have only one successor and, except *first*, one predecessor. In addition, the first activity of day d should be the same as the last activity of day $d-1$. Similarly, the first location of day d should be the same as the last location of day $d-1$. Regarding mode, the previous consistency constraints apply: the mode can change if the individual returns home between the end of the current day and the start of the new one.

2. We consider that the first activity ($d = 0$) of the *first* schedule should be *dawn*. Similarly, the last activity of the *last* schedule ($d = T$) should be *dusk*. For the

other days ($0 < d < T$), the first activity can either be *dawn* or the last activity of the previous day, and the last activity can either be *dusk* or any other activity (including travel).

3. The schedule of each day S_d is still constrained to the time budget of $T_d = 24h$. The overall time budget can therefore be written $T = nT_d$, where n is a strictly positive integer indicating the number of days to be scheduled.

Considering A_d the set of n_d activities of day d, with a^d_{first} the first activity of the day and a^d_{last} the last activity of the day, we can add the multiday consistency constraints (28)-(30) to the problem.

Constraints (8)-(25) defined for the single day problem still apply for each day d.

$$\omega^{d=0}_{\text{dawn}} = \omega^{d=T}_{\text{dusk}} = 1, \tag{28}$$

$$\omega^{d}_{a_{\text{first}}} = \omega^{d-1}_{a_{\text{last}}}, \qquad \forall 0 < d \leq T \tag{29}$$

$$z_{b,a_{\text{first}}} = z_{a_{\text{last}},b} = 0, \qquad \forall b \in A_d, \tag{30}$$

4 Empirical investigation

We test the modified OASIS framework to simulate multiday schedules for a sample of individuals from the MOBIS dataset (Molloy *et al.*, 2022). The MOBIS dataset is a longitudinal dataset conducted in Switzerland, and which contains 8 weeks of GPS traces for 3680 respondents. For the analyses presented in this paper, and for the sake of efficiency, we have randomly selected a subsample of 460 respondents.

In this section, we present first the implementation of the modelling hypotheses presented in Section 3.2. Then, we illustrate the output of the framework for a small problem. We discuss limitations and identify the axes that should be the focus of future research.

4.1 Preferences

One crucial hypothesis of the multiday analysis is that individuals' preferences for activity timings are specific to the day, meaning that people do not penalise equally schedule deviations from their preferences. The preferences might also be different depending on the type of day. The most intuitive categorisation of days is weekdays (Monday to Friday) versus weekends (Saturday to Sunday).

By analysing the dataset, we observe a behaviour that seems to confirm this assumption (Figure 2). For the *work* activity (Figure 2(a)), the distribution of start times shows that this activity is significantly more present in weekday schedules as opposed to weekends. In addition, the weekday distribution presents two peaks at 6:00 and at 11:00, whereas the weekend distribution is more uniform. The same observations can be made for *education*.

For *leisure* (Figure 2(c)), we see that the distributions for weekday and weekends are comparable in terms of frequency counts, with a higher peak in the evening at 17:00, and earlier in the day at 13:00 for weekends. Both sets present a significant peak at midnight, indicating that leisure is often scheduled overnight.

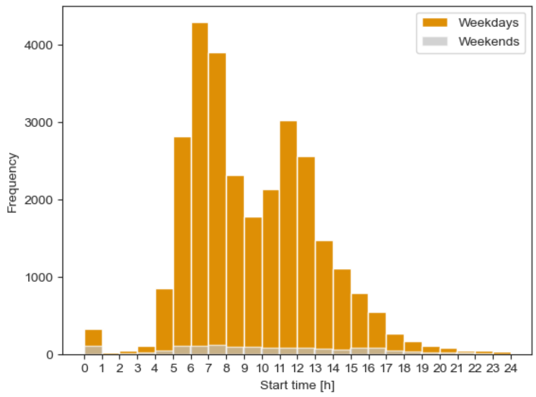
The conclusions for *shopping* (Figure 2(d)) are similar to leisure: on weekdays the distribution shows a peak in the second half of the day (between 14:00 and 17:00), while on weekends the start times are more evenly distributed during the day.

These observations support our assumptions: the preferences towards activities is different depending on the day. Some activities are scheduled more or less often, and the preferred start times and durations are different.

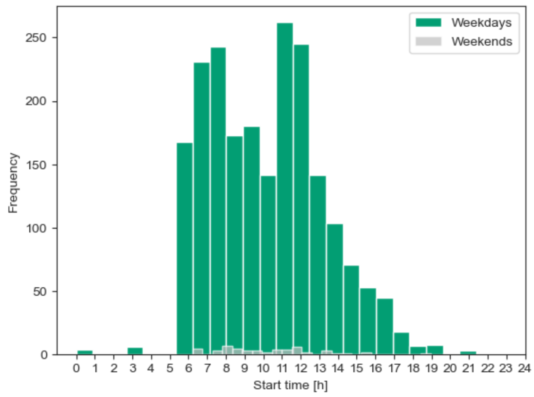
We integrate these activity and day-specific preferences for start time and duration to the model by fitting log-normal distributions to the observed schedules (start time, duration, and frequency) of the sample over 8 weeks. The parameters of these distributions are reported in Table 2. For each simulation and individual, we draw values from these distributions, that will be used as their preferred start time $x^*_{a,d}$, duration $\tau^*_{a,d}$ and activity frequency $f^*_{a,d}$.

Regarding activity participation, we observe notable trends for the weekly frequency (Figure 3) which is defined as the number of days in a week when a given activity is

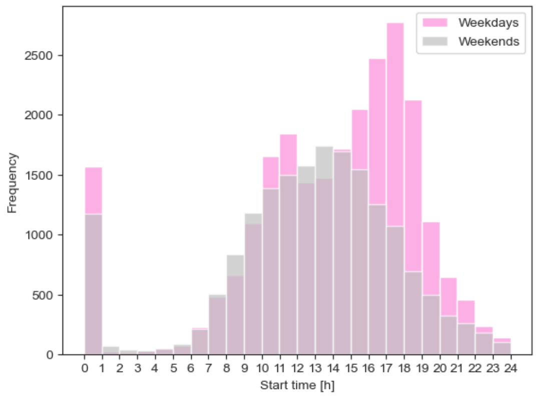
Figure 2: Distribution of start times (Mon-Fri vs. Sat-Sun) for different activities.



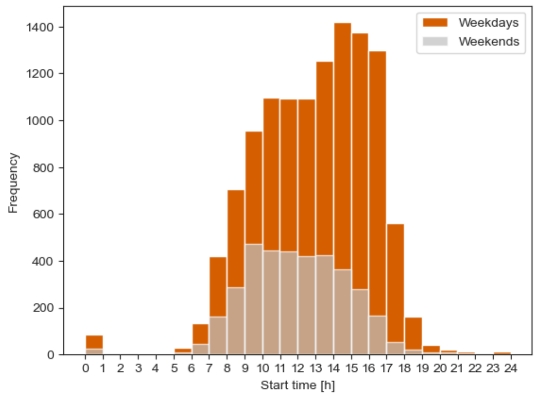
(a) Work



(b) Education



(c) Leisure



(d) Shopping

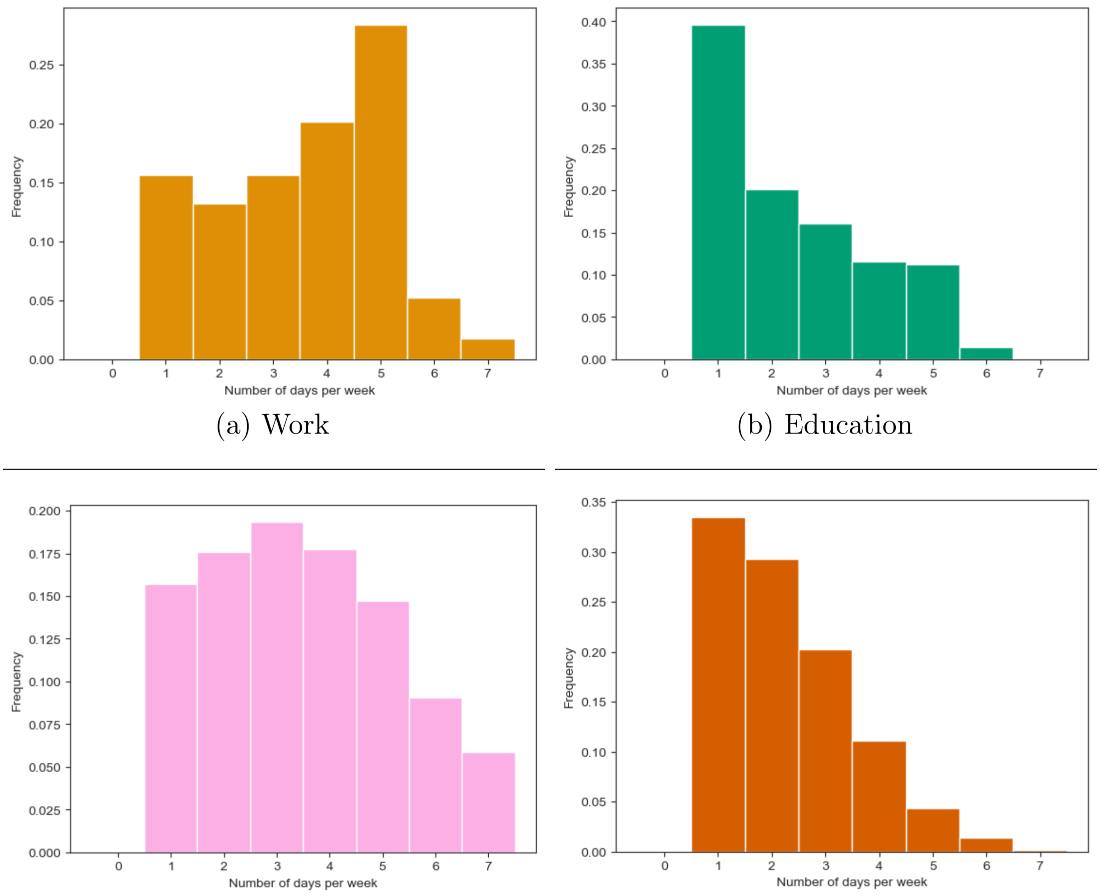
Table 2: Desired times distributions in sample

Day	Activity	Start time	Duration
	Home	Log-$\mathcal{N}$ $(8.7, -4.1, 2.8)$	Log-$\mathcal{N}$ $(0.34, -5.5, 10.3)$
	Work	Log-$\mathcal{N}$ $(0.15, -14.7, 24.0)$	Log-$\mathcal{N}$ $(1.4, 0, 1.6)$
Weekdays	Education	Log-$\mathcal{N}$ $(0.14, -11.8, 21.8)$	Log-$\mathcal{N}$ $(0.74, -0.51, 2.3)$
	Leisure	Log-$\mathcal{N}$ $(9.1, -4.8, 4.0)$	Log-$\mathcal{N}$ $(1.3, 0, 0.88)$
	Shopping	Log-$\mathcal{N}$ $(0.015, -204.1, 216.9)$	Log-$\mathcal{N}$ $(0.85, 0, 0.27)$
	Home	Log-$\mathcal{N}$ $(8.7, -4.1, 2.8)$	Log-$\mathcal{N}$ $(0.49, -3.9, 9.4)$
	Work	Log-$\mathcal{N}$ $(0.06, -81.1, 91.3)$	Log-$\mathcal{N}$ $(1.7, 0.06, 1.4)$
Weekends	Education	Log-$\mathcal{N}$ $(0.35, 2.9, 6.9)$	Log-$\mathcal{N}$ $(1.6, 0.06, 1.4)$
	Leisure	Log-$\mathcal{N}$ $(8.8, -6.8, 4.1)$	Log-$\mathcal{N}$ $(1.3, 0, 0.68)$
	Shopping	Log-$\mathcal{N}$ $(0.012, -244.3, 256.2)$	Log-$\mathcal{N}$ $(0.90, 0, 0.30)$

scheduled at least once, and the daily frequency (Figure 4), which is the number of time an activity is performed in a given day.

For weekly participation, we notice that *work* (Fig. 3(a)) is a very regular activity, with a majority of individuals working for 4 days or more per week. *Education* (Fig. 3(b)) is scheduled less frequently (mostly once per week), but it is also regular to the extent that this peak is significant as compared to the other options. This indicates a fairly low variability across individuals and weeks. On the other hand, *shopping* but especially *leisure* have less defined peaks, and the participation remains the same across a wider range (from once to 5 times per week for leisure, and from once to three times per week for shopping). Contrasted with work and education, these activities seem more occasional, or irregular in terms of weekly patterns.

Figure 3: Distribution of weekly participation frequency for different activities.



(a) Work

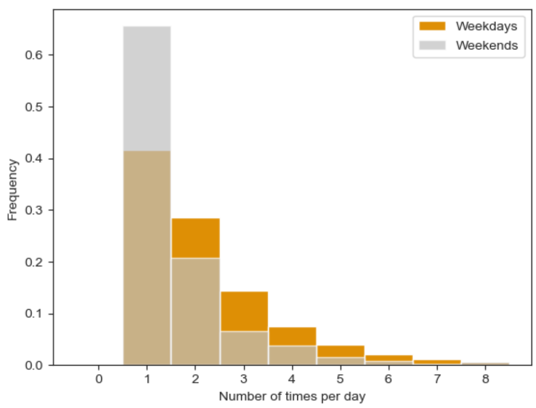
(b) Education

(c) Leisure

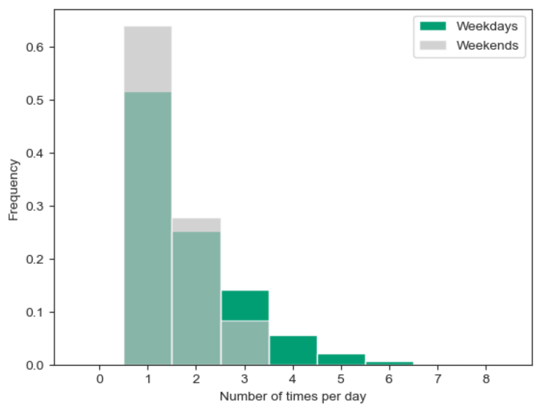
(d) Shopping

The distributions of daily participation frequency are very similar across activities and across days. For *work*, *education* and *shopping* (Figs. 4(a), 4(b) and 4(d)) the majority of activities are only performed once per day both for weekdays and weekend. As expected, *leisure* (Fig. 4(c)) is the activity that is the most frequently performed multiple times in a day, especially on weekends.

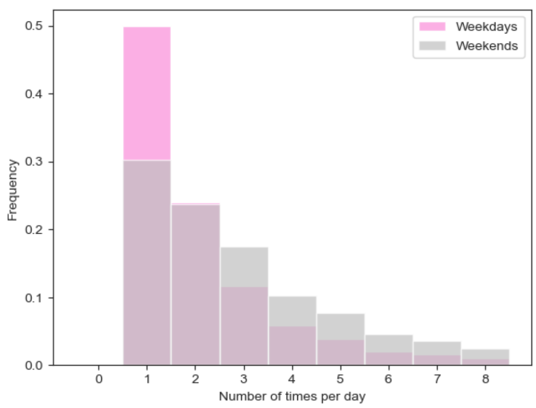
Figure 4: Distribution of daily participation frequency (Mon-Fri vs. Sat-Sun) for different activities.



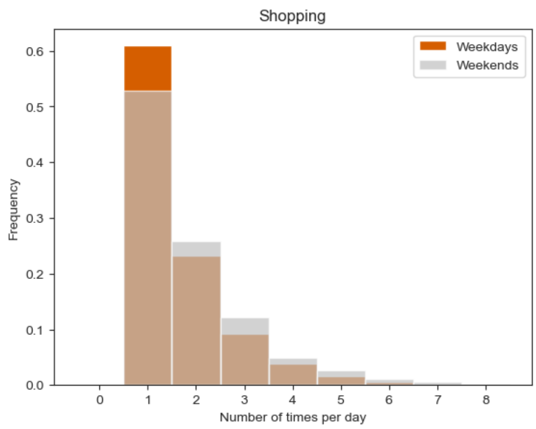
(a) Work



(b) Education



(c) Leisure



(d) Shopping

The results of Figures 3 and 4 can be used to define the frequency preferences for each individual. In this case, we use the mean of the distributions to approximate these values. They are reported in Table 3.

Table 3: Desired weekly and daily activity frequencies

Activity	Weekly frequency	Daily frequency
Home	4.9	2.6
Work	3.5	2.2
Education	2.4	1.8
Leisure	3.5	2.6
Shopping	2.3	1.7

4.2 Weights

As described in the previous section, the weights of the multiobjective function should translate a notion of priority between days. In this study, we investigate the impact of the weights on the optimisation results by testing two configurations:

1. There is no priority between days, and $w_d = 1\ \forall d$
2. The scheduling process prioritises weekdays over weekends. Weekdays therefore have a higher weight, such that the multiobjective function can be written:

$$\Omega = w_{\text{weekday}} \sum_i U_S^i + w_{\text{weekend}} \sum_j U_S^j \qquad \forall i \in D_{\text{weekdays}}, \forall j \in D_{\text{weekend}}$$

with D_{weekdays} the set of weekdays (Monday to Friday), and D_{weekend} the set of weekend days (Saturday and Sunday). For this second scenario, we arbitrarily define: $w_{\text{weekday}} = 0.8$ and $w_{\text{weekend}} = 0.2$.

4.3 Utility parameters

For the parameters of the daily utility functions, we use the default OASIS activity-specific parameters (Pougala *et al.*, 2022). They are summarised in Table 4.

Table 4: OASIS activity-specific model parameters

	Parameter	Param. estimate	Rob. std err	Rob. t-stat	Rob. p-value
1	$\gamma_{\text{education}}$	18.7	3.17	5.89	3.79e-09
2	$\theta^{\text{early}}_{\text{education}}$	-1.35	0.449	-3.01	0.00264
3	$\theta^{\text{late}}_{\text{education}}$	-1.63	0.416	-3.91	9.05e-05
4	$\theta^{\text{long}}_{\text{education}}$	-1.14	0.398	-2.86	0.00428
5	$\theta^{\text{short}}_{\text{education}}$	-1.75	0.457	-3.84	0.000123
6	γ_{leisure}	8.74	1.94	4.50	6.79e-06
7	$\theta^{\text{early}}_{\text{leisure}}$	-0.0996	0.119	-0.836	0.403*
8	$\theta^{\text{late}}_{\text{leisure}}$	-0.239	0.115	-2.07	0.0385
9	γ_{shopping}	10.5	2.20	4.78	1.74e-06
10	$\theta^{\text{early}}_{\text{shopping}}$	-1.01	0.287	-3.51	0.000443
11	$\theta^{\text{late}}_{\text{shopping}}$	-0.858	0.237	-3.63	0.000284
12	γ_{work}	13.1	2.64	4.96	7.16e-07
13	$\theta^{\text{early}}_{\text{work}}$	-0.619	0.217	-2.85	0.00438
14	$\theta^{\text{late}}_{\text{work}}$	-0.338	0.168	-2.02	0.0438
15	$\theta^{\text{long}}_{\text{work}}$	-1.22	0.348	-3.51	0.000441
16	$\theta^{\text{short}}_{\text{work}}$	-0.932	0.213	-4.37	1.23e-05

4.4 Example

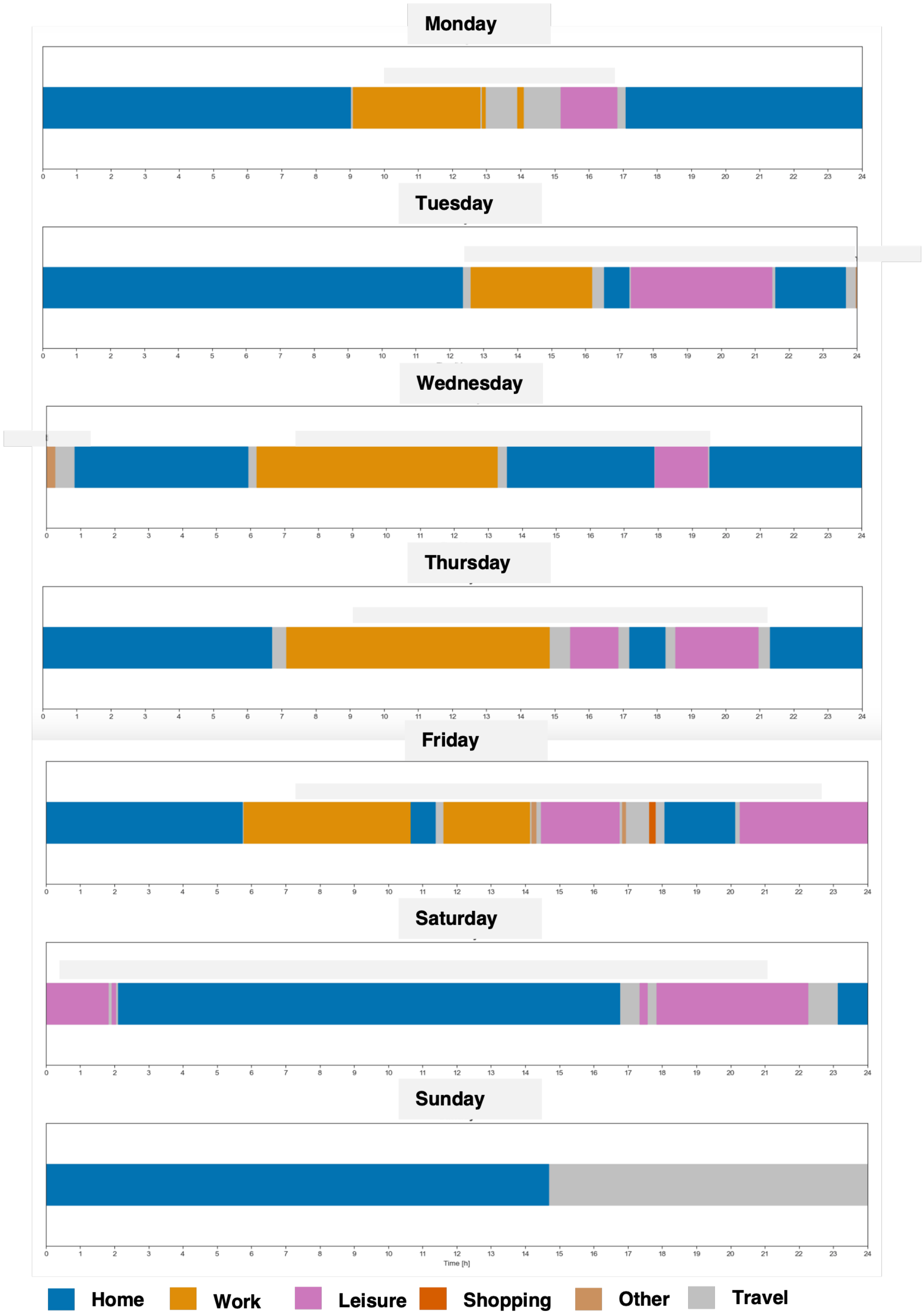
We solve a week of activities for an individual of the MOBIS dataset. The schedules are illustrated in Figure 5. This individual goes to work every weekday, with at least one leisure activity in the evening. They go shopping on Friday, and have two overnight activities between Tuesday and Wednesday, and between Friday and Saturday.

We use this example to test the multiday specification and methodology described in Section 3.2. For this individual, we provide as input the set of activities home, work, leisure, shopping, other, to be performed any day during the week, with the penalties and preferences introduced in Sections 4.1 and 4.3.

We compare the results of the multiday model with the outputs of the single day model that we run n_{days} times, to investigate the added value of our methodological extension. This procedure should be regarded as a proof of concept, and not as a formal validation. We discuss in this section future work required to evaluate the contribution more formally.

Table 5 summarises the tested scenarios:

Figure 5: Example week from MOBIS dataset



1. Scenario 1: We generate 2 consecutive weekdays. We do not consider any hierarchy between days, and therefore the multiobjective function is not weighted.

2. Scenario 2: We generate one weekend (Saturday and Sunday). We do not consider any hierarchy between days, and therefore the multiobjective function is not weighted.

3. Scenario 3: We generate a Friday and a Saturday. as the days belong to different categories, we compare the results of the model with the non-weighted and weighted objective functions.

Table 5: Simulation scenarii

Model	Benchmark	Unweighted multiday	Weighted multiday
Scenario 1 (weekdays)	✓	✓	
Scenario 2 (weekends)	✓	✓	
Scenario 3 (Friday-Saturday))	✓	✓	✓

4.4.1 Results

Figure 6 to Figure 11 show selected schedules that were generated from the models for the scenarios presented in Table 5.

Using the single day OASIS model (Figs. 6 and 9) generally leads to schedules that resemble the observed activity patterns, especially regarding the work activity. However, the specification of the model does not allow for overnight activities, and generates significantly more full days at home than the other models and the observed week. For Scenario 2, almost all simulated schedules were spent at home.

For both multiday specifications, we observe some overnight activities (significantly more than in the observed data), and they generate fewer days fully at home. We notice that the activity durations do not match the preferences (e.g. leisure and other), but more importantly that there does not seem to be any distinction between days, especially for the unweigthed model. For example, as seen in Figure 8, work is often scheduled on Sundays, which does not reflect the majority of schedules in the dataset. On the other hand, the inclusion of weights in the multiobjective utility function seems to have an impact: in Figure 11, we see that, while there is some variety in the activities scheduled on Friday, Saturday does not vary much.

These preliminary results show the potential of the multiday analysis as opposed to simulating independent single days, but also highlights some crucial limitations to be addressed in future work:

- There needs to be a higher differentiation between days. This can either be done through the utility function or the constraints. Regarding the utility function, the issue stems from the fact that we have only used one set of parameters for every simulation. It is therefore important to estimate different parameters per day or category of day. For the constraints, some context-dependent rules might aid the realism of the results (e.g. working hours, opening days of shops...).

- In addition to frequency, there needs to be a corrective or penalising term in the utility function to limit the length or extent of the overnight activities. In this case, as we have not constrained the return home or estimated preferred timings or duration at home, the utility for performing certain activities at unrealistic times is too high.

5 Conclusion and future research

This paper presents a preliminary investigation on the extension of the single-day integrated activity scheduler with intrapersonal interactions in the form of behaviour over multiple days. We have started by adapting the specification of the optimisation problem by including multiday-specific variables and parameters to the utility function, and multiday constraints. Notably, we have found that relaxing some constraints is necessary to allow for specific observations such as overnight activities.

In future work, we will improve the methodology by focusing on the following points.

Figure 6: Example schedules for Scenario 1 (Single day)

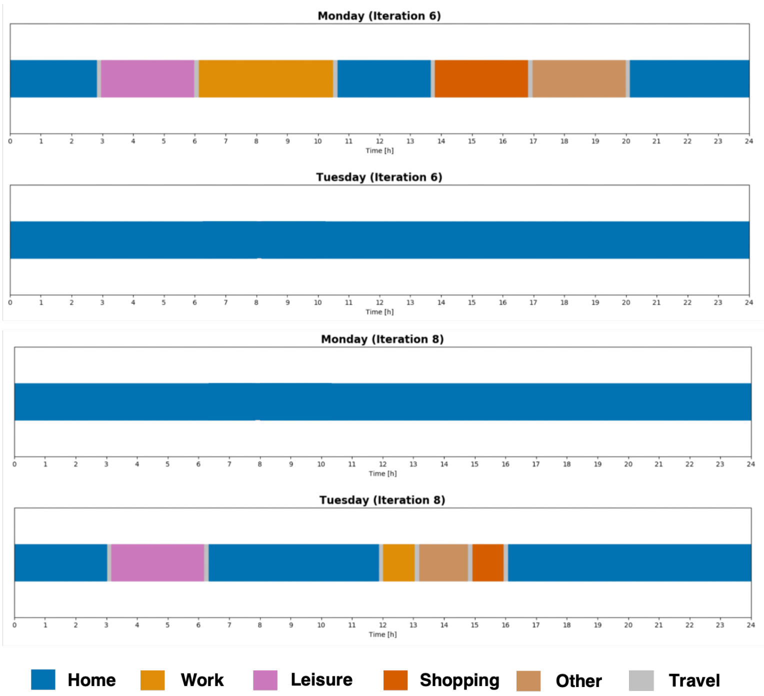


Figure 7: Example schedules for Scenario 1 (Multiday unweighted)

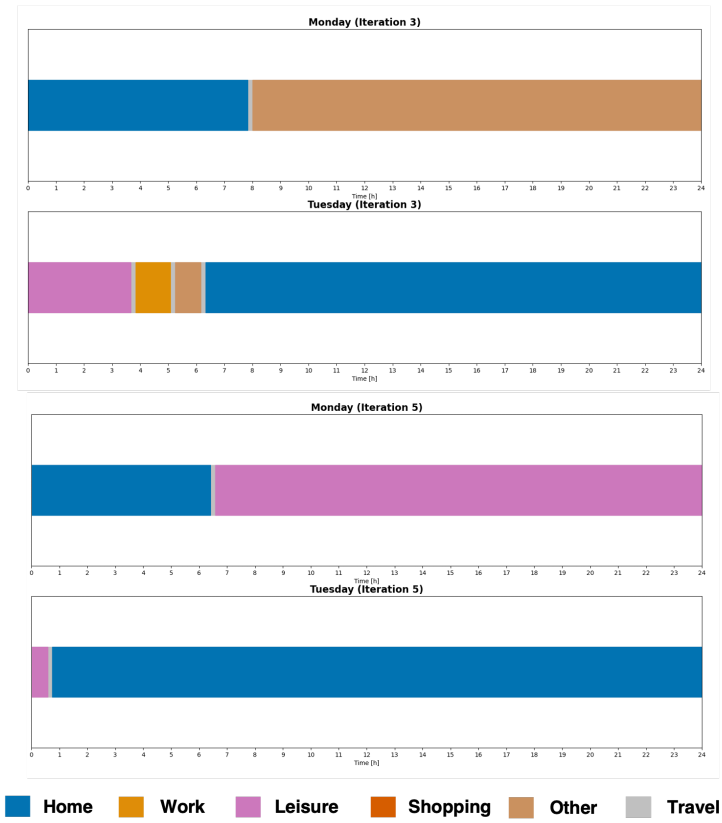
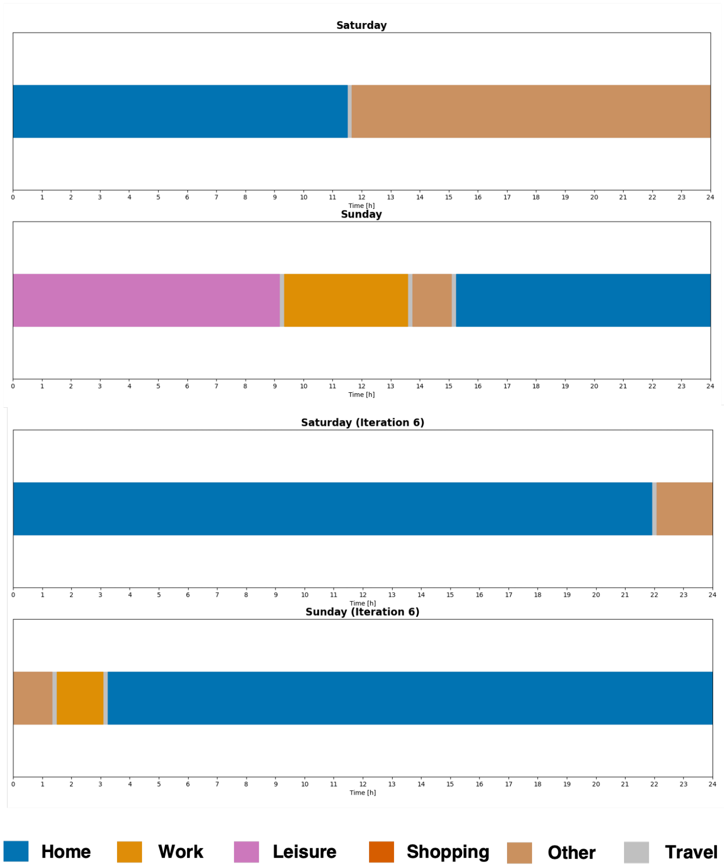


Figure 8: Example schedules for Scenario 2 (Multiday unweighted)



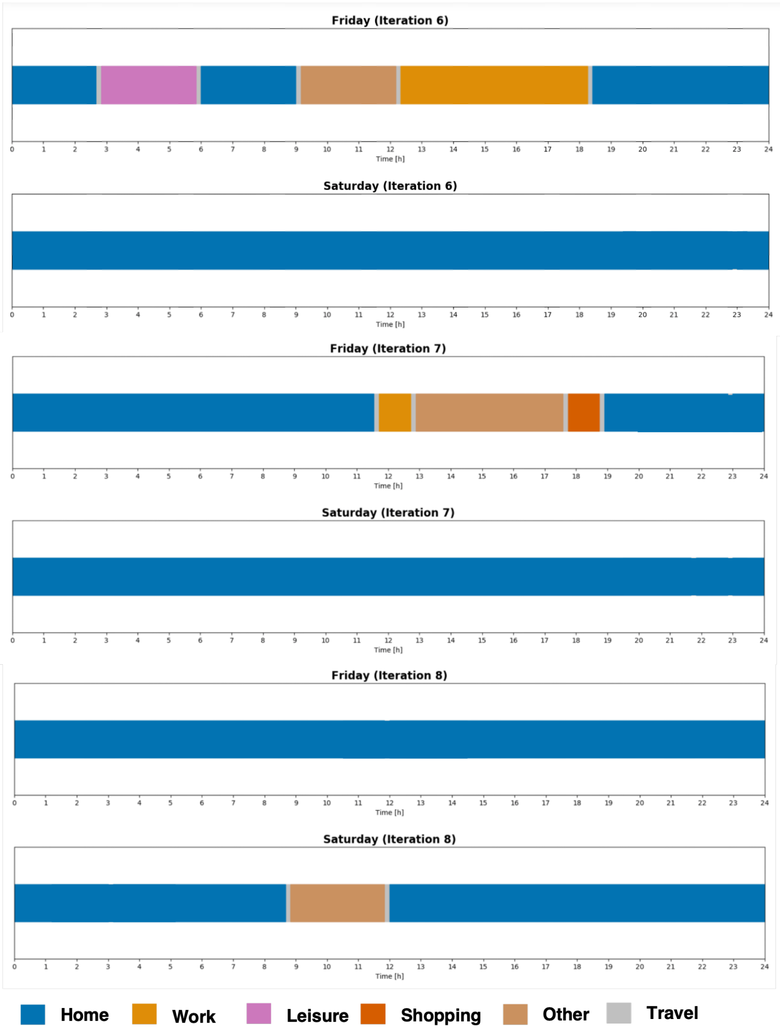
- We will test appropriate utility specifications to better accommodate correlation, to capture both habits (which we have started investigating in this paper), and forward-looking behaviour. This could be achieve by including individual expectations for future activity or schedule utilities in the current function.

- We will investigate how to adapt the single-day parameter estimation methodology of OASIS. Some questions must be addressed, such as the existence of a different choice set for each day or category of days (e.g., weekdays vs. weekends)

- Validation will be an important research focus; specifically, developing indicators to evaluate the simulation performance over longer periods of time. One initial step would be to use aggregated insights such as activity motifs or patterns (Schultheiss, 2021), which can be compared to the simulated motifs.

- A dedicated case study will be developed, in order to perform sensitivity analyses, and identify computational or efficiency limits of the framework (e.g. runtime as a

Figure 9: Example schedules for Scenario 3 (Single day)



function of the number of days to schedule.) This point is especially important to test the scalability of the OASIS framework to large-scale applications.

6 References

Arentze, T. A., D. Ettema and H. J. Timmermans (2011) Estimating a model of dynamic activity generation based on one-day observations: Method and results, *Transportation Research Part B: Methodological*, **45** (2) 447–460, feb 2011, ISSN 01912615.

Bierlaire, M., E. Frejinger and T. Hillel (2021) Dynamic choice models, *Technical Report*,

Figure 10: Example schedules for Scenario 3 (Multiday unweighted)

TRANSP-OR 210305, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

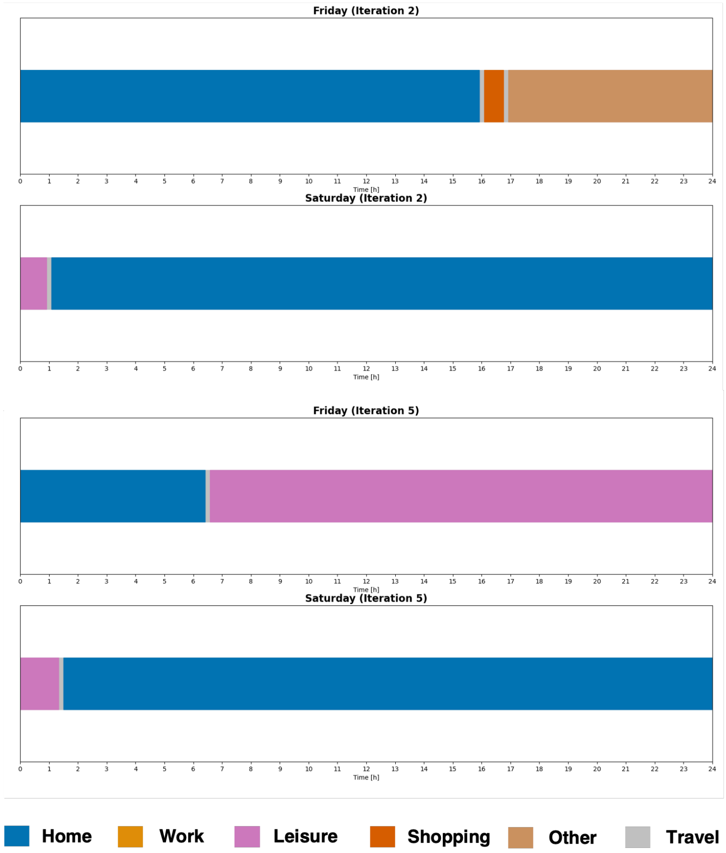
Calastri, C., S. Hess, A. R. Pinjari and A. Daly (2020) Accommodating correlation across days in multiple-discrete continuous models for time use, *Transportmetrica B*, **8** (1) 90–110, jan 2020, ISSN 21680582.

Hilgert, T., M. Heilig, M. Kagerbauer and P. Vortisch (2017) Modeling Week Activity Schedules for Travel Demand Models, *Transportation Research Record: Journal of the Transportation Research Board*, **2666** (1) 69–77, jan 2017, ISSN 0361-1981.

Molloy, J., A. Castro, T. Götschi, B. Schoeman, C. Tchervenkov, U. Tomic, B. Hintermann and K. W. Axhausen (2022) The MOBIS dataset: a large GPS dataset of mobility behaviour in Switzerland, *Transportation*, ISSN 1572-9435.

Nurul Habib, K. M. and E. J. Miller (2008) Modelling daily activity program generation considering within-day and day-to-day dynamics in activity-travel behaviour,

Figure 11: Example schedules for Scenario 3 (Multiday weighted)



Transportation, **35** (4) 467–484, jul 2008, ISSN 00494488.

Pougala, J., T. Hillel and M. Bierlaire (2022) Capturing trade-offs between daily scheduling choices, *Journal of Choice Modelling*, **43**, 100354, jun 2022, ISSN 1755-5345.

Roorda, M. J. and T. Ruiz (2008) Long- and short-term dynamics in activity scheduling: A structural equations approach, *Transportation Research Part A: Policy and Practice*, **42** (3) 545–562, ISSN 0965-8564.

Schultheiss, M.-E. (2021) Exploring causalities between modal habits, activity scheduling, and multi-day locational practices.

Zhang, W., C. Ji, H. Yu, Y. Zhao and Y. Chai (2021) Interpersonal and intrapersonal variabilities in daily activity-travel patterns: A networked spatiotemporal analysis, *ISPRS International Journal of Geo-Information*, **10** (3), ISSN 2220-9964.