Specification, estimation and validation of a pedestrian walking behavior model

Gianluca Antonini

Javier Cruz

Michel Bierlaire

Thomas Robin

14th March 2008

Motivation

- Model the pedestrian behaviour

Motivation

- Model of the pedestrian behaviour

  ⇨ **incorporate the model in more global models**

- Use of econometric models

  ⇨ **visibility and control of the specification**

- Calibration on real data

  ⇨ **estimated parameters values**

State of the art

pedestrian modelling

limited estimation data set

huge estimation data set

use of machine learning methods

aggregate level

Flow problems

disaggregate behaviour

discrete models

- flexible
- lot of parameters

continuous models

- postulate equations
- few number of parameters

Objectives

- Model the pedestrian behaviour at **operational** level
- Develop a specification with **‘constrained’** and **‘unconstrained’** parameters
- **Estimate** the model
- **Validate** the model
- Implement the model in a **simulator**

Outline

- Introduction
- Discrete choice models
- Model specification
- Model estimation
- Model validation
- Simulator
- Conclusion

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Introduction

- **Microscopic** model : capture the behavior of **each** pedestrian

  → **Discrete choice model**

- Different **behavioral levels** :

  Strategical : destination

  Tactical : route choice

  Fixed

  Operational level : short range behavior

  instantaneous decisions

- Concept of **personal space** : interactions with other pedestrians

  Leader follower

  Collision avoidance

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

- Introduction
- **Discrete choice models**
- Model specification
- Model estimation
- Model validation
- Simulator
- Conclusion

Discrete choice models : introduction

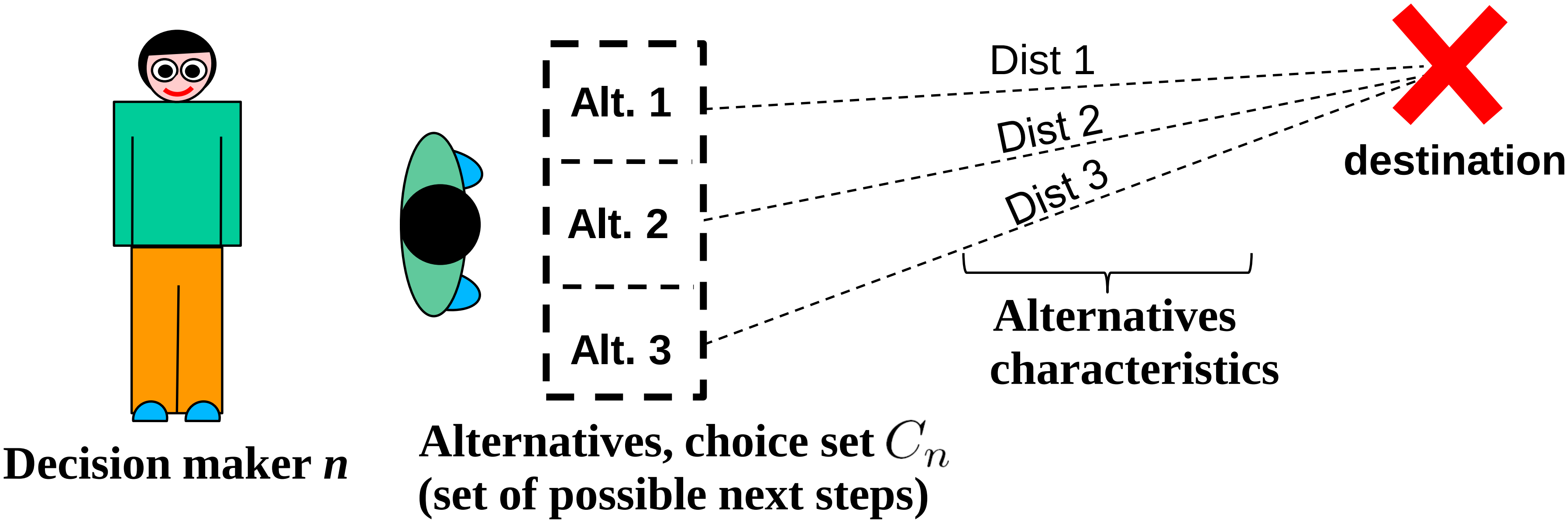
- **Econometric** models developped since the 50's
- **Disagregate** model
- A choice theory defines :
  - **a decision maker** : each pedestrian
  - **alternatives** : possible immediate future steps
  - **attributes of alternatives** : characteristics
  - **decision rule :** utility maximisation theory

TRANSP-OR

EPFL
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Discrete choice model : our context

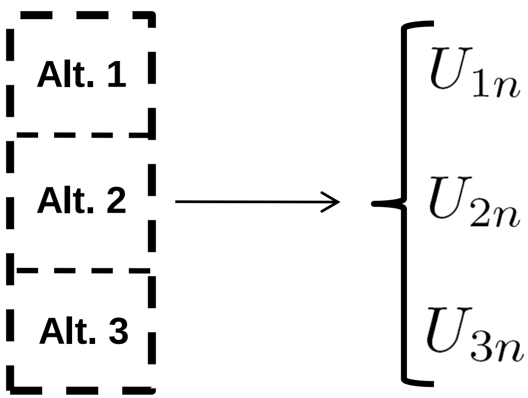
- At **each step** the pedestrian has to choose the next step in a choice set
- **Example :** only considering distance toward destination



Which alternative will he choose?

Discrete choice model : decision rule

- Utility maximisation theory
- Association of a function, called **utility** to each alternative
- It depends on the **alternative *i*** , and on the **decision maker *n***



The decision maker ***n*** will choose the alternative ***i*** which has the **higher utility**

Discrete choice model : utility function

- Utility is a **latent concept**
- It can not be directely observed
- Decision maker: stochastic decision rules
- Analyst: Lack of information

Uncertainty

$$U_{in} = V_{in} + \epsilon_{in}$$

V_{in} **: deterministic part of the utility of alternative *i* for individual *n***

ϵ_{in} **: error term, different assumptions can be made on its distribution**

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Discrete choice model : utility specification

- **Example** : V_{in} depends only on the distance toward the destination

$$V_{in} = \beta dist_i$$

β : unknown parameter, **has to be estimated from the data**

$dist_i$ **:** distance between alternative ***i*** and the final destination

→ **How do we estimate β ?**

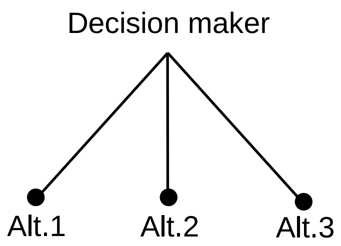
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Discrete choice model : error term

- **Example** : suppose ϵ_{in} independent and identically distributed (iid) with an extreme value distribution

⟹ **Multinomial Logit model**



$P_n(i|C_n)$: probability for the individual ***n*** to choose the alternative ***i***

$$P_n(i|C_n) = \frac{e^{V_{in}}}{\sum_{j \in C_n} e^{V_{jn}}}$$

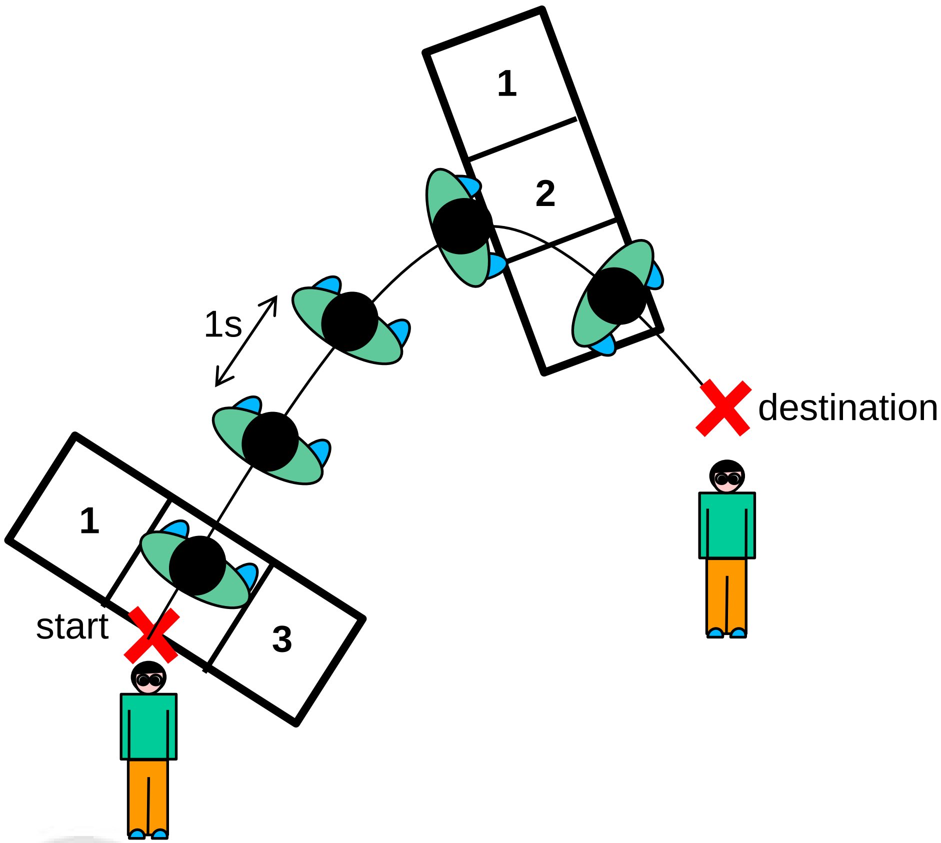
C_n : Choice set, depends on the individual ***n***

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Discrete choice model : estimation data

Pedestrian trajectories



Data set

Observation	Choice	dist.1	dist.2	dist.3
1	2			
2	2			
3	2			
4	2			
5	3			
6	2			

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Discrete choice model : likelihood

- Maximisation of the likelihood function

$$l(\beta) = \prod_{n \in N} \prod_{i \in C_n} (P_n(i|C_n))^{y_{in}} = \prod_{n \in N} \prod_{i \in C_n} (\frac{e^{\beta dist_i}}{\sum_{j \in C_n} e^{\beta dist_j}})^{y_{in}}$$

y_{in} : indicator equals to 1, if individual ***n*** has choosen alternative ***i***, 0 otherwise

N : set of individuals in the population

- In practice use of the log-likelihood (numerical reasons)

$$L(\beta) = \sum_{n \in N} \sum_{i \in C_n} y_{in} log(\frac{e^{\beta dist_i}}{\sum_{j \in C_n} e^{\beta dist_j}})$$

$\Rightarrow$ β **estimated from the data**

Why using a discrete choice model?

- **Disaggregate**

  → capture the behaviour of each pedestrian

- **Flexible**

  → easy to add behavioural modules

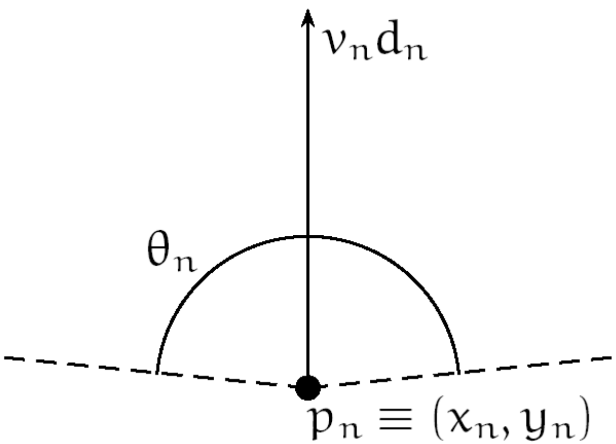
  → easy to add pedestrians characteristics

- **Estimation**

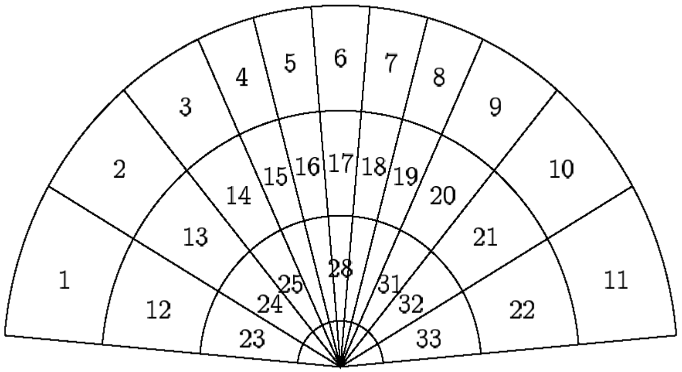
  → estimation on a real data set by likelihood maximisation

- Introduction
- Discrete choice models
- **Model specification**
- Model estimation
- Model validation
- Simulator
- Conclusion

Model specification : the space discretization



Pedestrian visual space



Choice set : discretization of the visual space

At each step the **choice set** depends on the pedestrian **speed** and **direction**

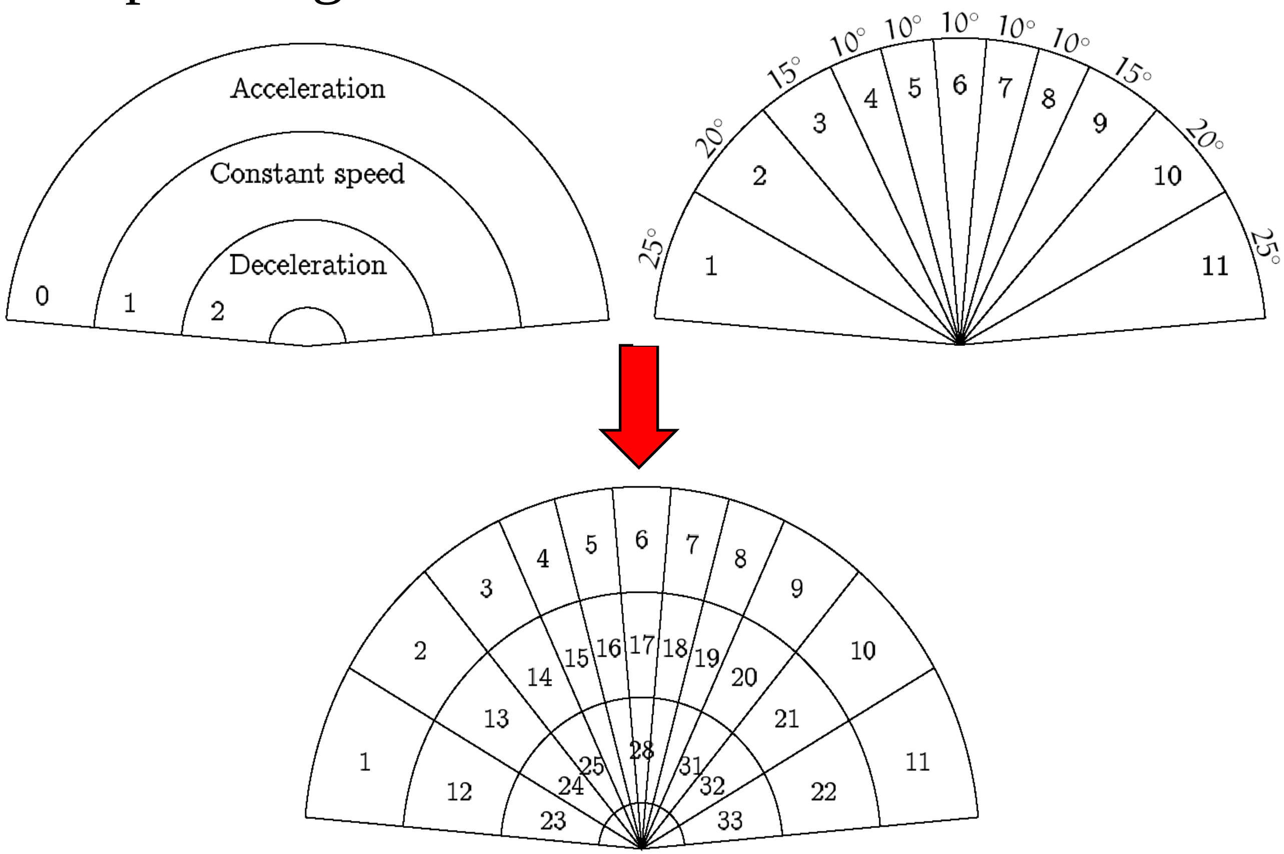
Dynamic choice set

TRANSP-OR

EPFL ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model specification : the choice set

3 speed regimes | 11 directions



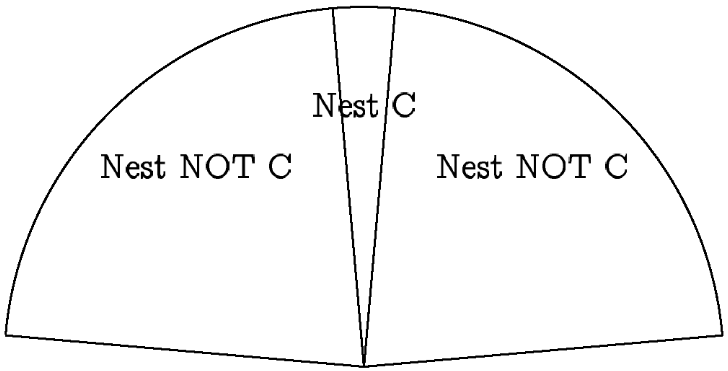
33 alternatives

Model specification : cross nested structure

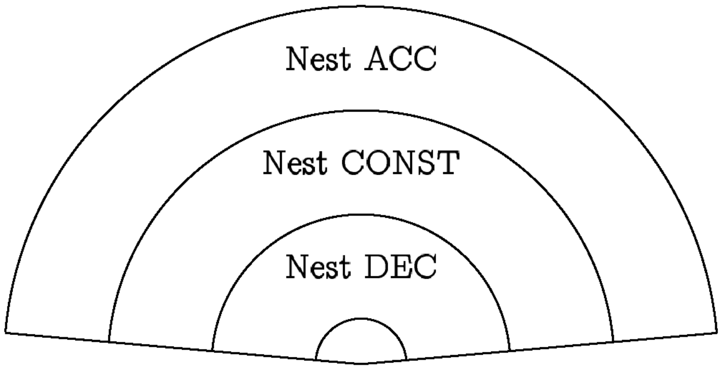
- **Hypothesis** : alternatives correlated along speed regimes and directions

  ➡ **Cross Nested Logit model**

- **Cross Nested structure** : each alternative belongs to 2 nests

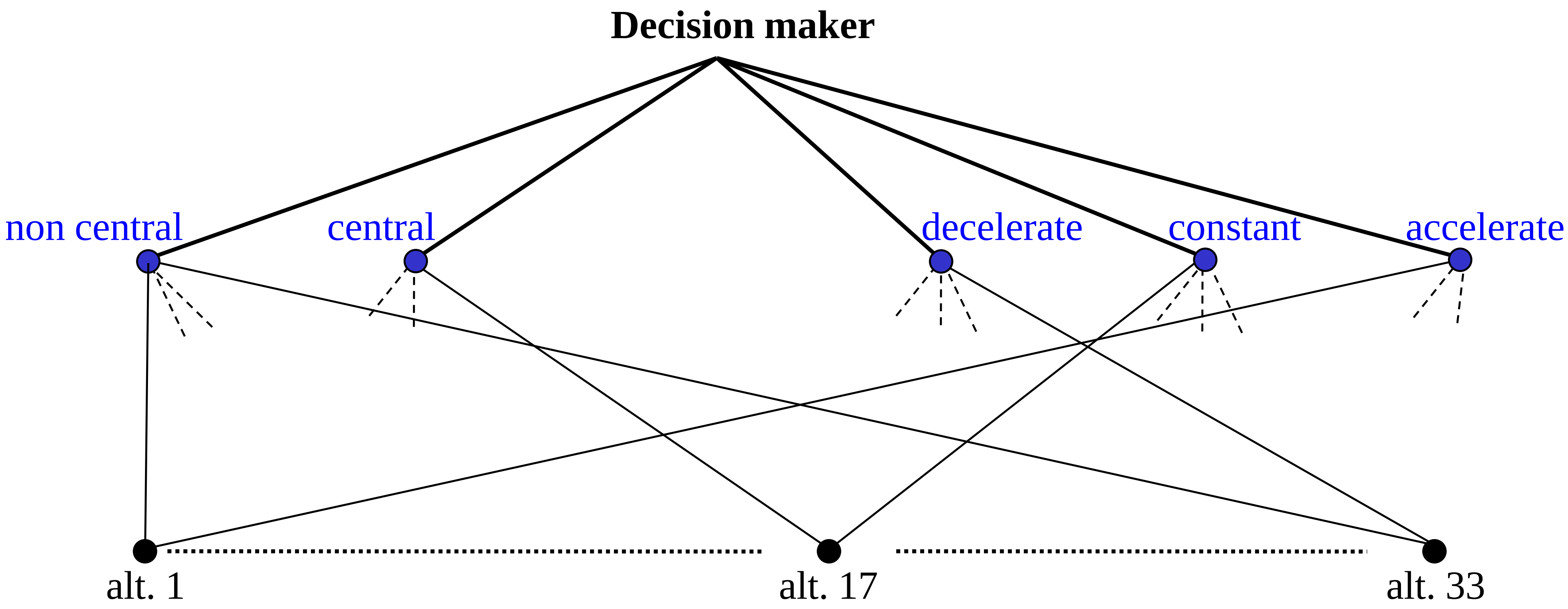


Nesting based on direction



Nesting based on speed regime

Model specification : cross nested structure



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model specification : cross nested structure

- Probability of choosing the alternative i :

$$P(i|C) = \sum_{m=1}^{M} \frac{\left(\sum_{j\in C} \alpha_{jm}^{\mu_m/\mu} y_j^{\mu_m}\right)^{\frac{\mu}{\mu_m}}}{\sum_{n=1}^{M} \left(\sum_{j\in C} \alpha_{jn}^{\mu_n/\mu} y_j^{\mu_n}\right)^{\frac{\mu}{\mu_n}}} \frac{\alpha_{im}^{\mu_m/\mu} y_i^{\mu_m}}{\sum_{j\in C} \alpha_{jm}^{\mu_m/\mu} y_j^{\mu_m}}$$

C : choice set

M : number of nests

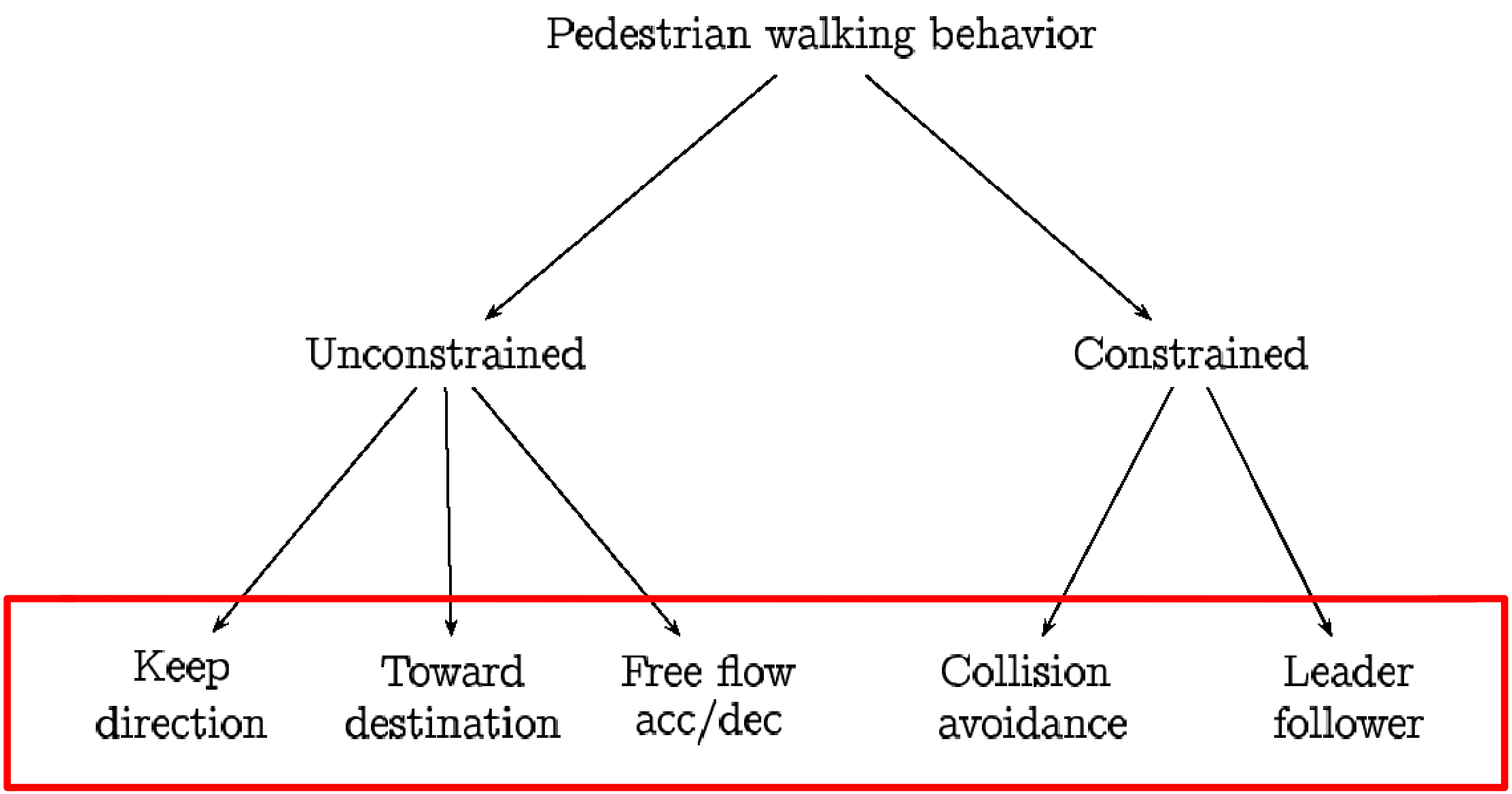
V_i : utility of alternative i

α_{jm} : membership degree of alternative j in the nest n

μ_m : parameter of the nest m

$y_i = e^{V_i}$

Model specification : utility specification



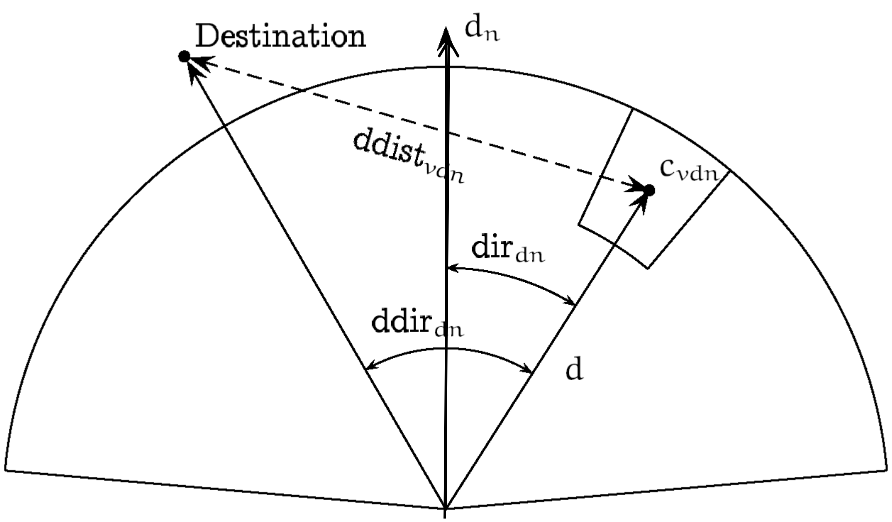
Model specification : utility specification

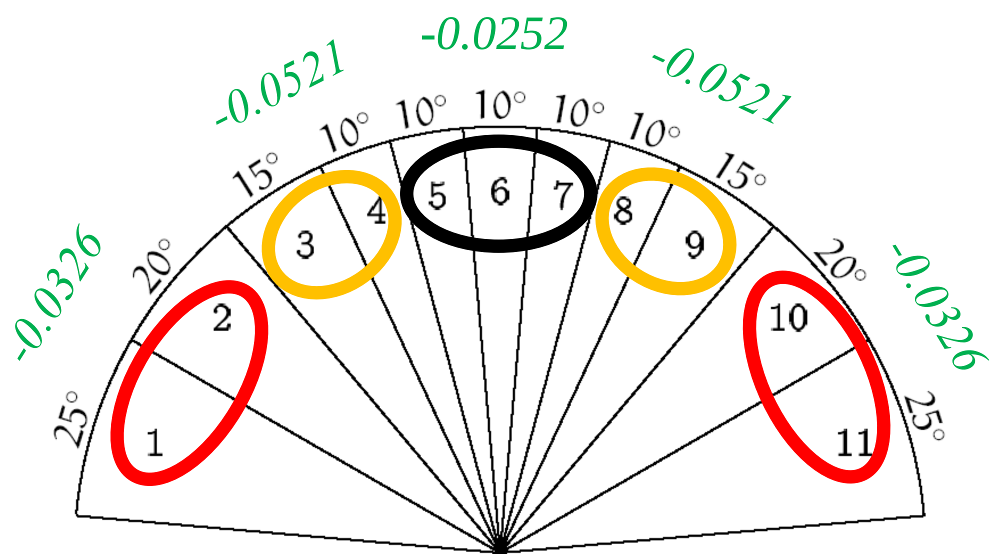
$$
\begin{aligned}
V_{vdn} = & \left.\begin{array}{r}
\beta_{dir\_central}\mathrm{dir}_{dn}I_{central} + \\
\beta_{dir\_side}\mathrm{dir}_{dn}I_{side} + \\
\beta_{dir\_extreme}\mathrm{dir}_{dn}I_{extreme} +
\end{array}\right\} \quad \textit{keep direction} \\
& \left.\begin{array}{r}
\beta_{ddist}\mathrm{ddist}_{vdn} + \\
\beta_{ddir}\mathrm{ddir}_{dn} +
\end{array}\right\} \quad \textit{toward destination} \\
& \left.\begin{array}{r}
\beta_{dec}I_{v,dec}(v_n/v_{max})^{\lambda_{dec}} + \\
\beta_{accLS}I_{LS}I_{v,acc}(v_n/v_{maxLS})^{\lambda_{accLS}} + \\
\beta_{accHS}I_{HS}I_{v,acc}(v_n/v_{max})^{\lambda_{accHS}} +
\end{array}\right\} \quad \textit{free flow acceleration} \\
& \left.\begin{array}{r}
I_{v,acc}I^{L}_{acc}\alpha^{L}_{acc}D_{L}^{\rho^{L}_{acc}}\Delta v_{L}^{\gamma^{L}_{acc}}\Delta\theta_{L}^{\delta^{L}_{acc}} + \\
I_{v,dec}I^{L}_{dec}\alpha^{L}_{dec}D_{L}^{\rho^{L}_{dec}}\Delta v_{L}^{\gamma^{L}_{dec}}\Delta\theta_{L}^{\delta^{L}_{dec}} +
\end{array}\right\} \quad \textit{leader-follower} \\
& \left.\begin{array}{r}
I_{d,d_n}I_{C}\alpha_{C}e^{-\rho_C D_C}\Delta v_{C}^{\gamma_C}\Delta\theta_{C}^{\delta_C}
\end{array}\right\} \quad \textit{collision avoidance}
\end{aligned}
$$

Model specification : utility specification

- Keep direction (**unconstrained**) :

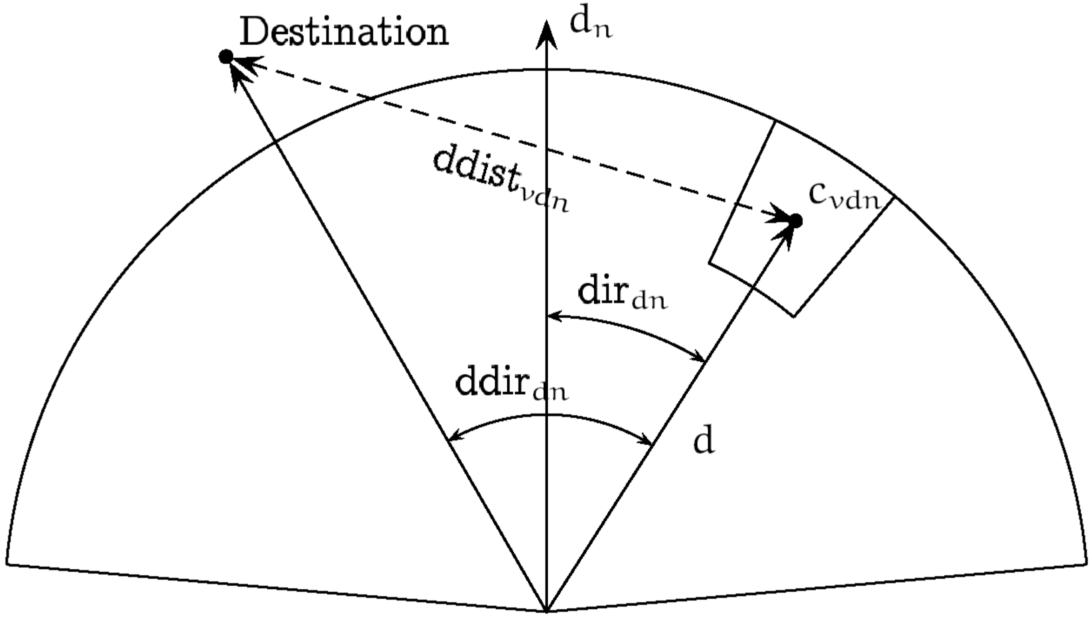
$$\underbrace{\overset{<0}{\beta_{\text{dir\_central}}}\,\text{dir}_{dn}\,I_{\text{central}}} + \underbrace{\overset{<0}{\beta_{\text{dir\_side}}}\,\text{dir}_{dn}\,I_{\text{side}}} + \underbrace{\overset{<0}{\beta_{\text{dir\_extreme}}}\,\text{dir}_{dn}\,I_{\text{extreme}}}$$





Model specification : utility specification

- Toward destination (**unconstrained**) : $\overset{\overset{-1.55}{<0}}{\beta_{\mathsf{ddist}}}\underbrace{\mathrm{ddist}_{\nu dn}}_{\textbf{distance}} + \overset{\overset{-0.0790}{<0}}{\beta_{\mathsf{ddir}}}\underbrace{\mathrm{ddir}_{dn}}_{\textbf{direction}}$

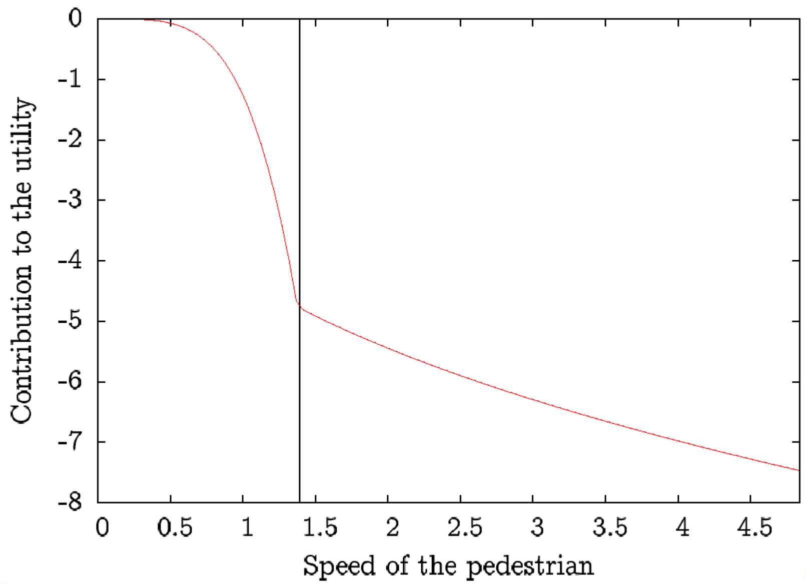


Model specification : utility specification

- Free flow acceleration (**unconstrained**) :

  \- Acceleration :

$$\underbrace{\overset{-4.97}{\beta_{accLS}} I_{LS} I_{v,acc} (v_n / v_{maxLS})^{\overset{4.16}{\lambda_{accLS}}}}_{\text{Low speed}} + \underbrace{\overset{-7.47}{\beta_{accHS}} I_{HS} I_{v,acc} (v_n / v_{max})^{\overset{0.358}{\lambda_{accHS}}}}_{\text{High speed}}$$



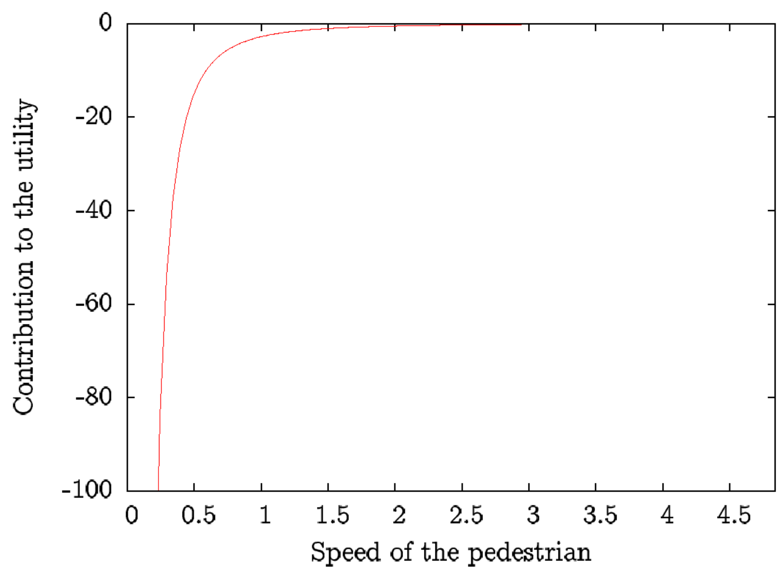
Model specification : utility specification

- Free flow acceleration (**unconstrained**) :

  - Deceleration :

-0.0630 *-2.41*

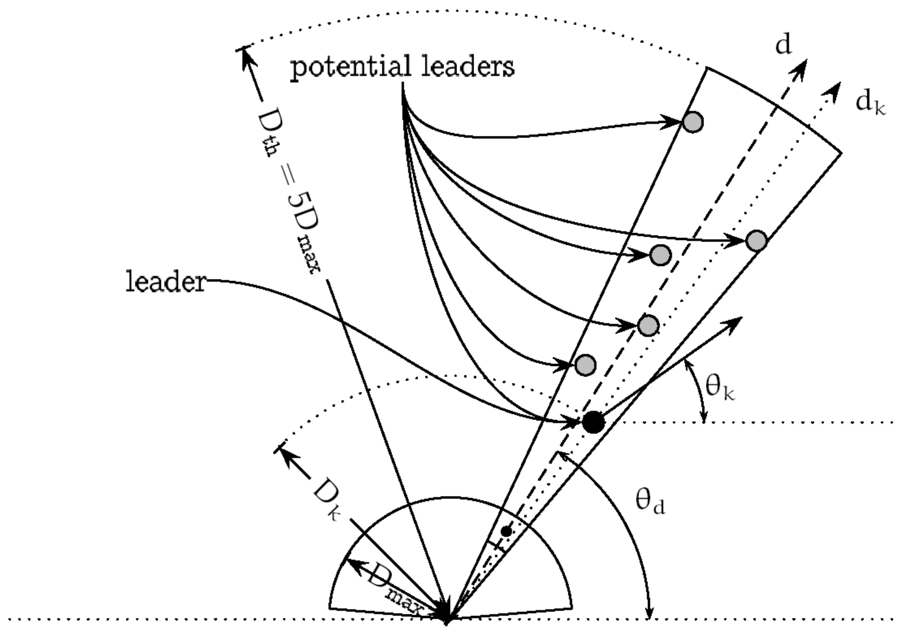
$$\beta_{dec} I_{v,dec} (v_n / v_{max})^{\lambda_{dec}}$$



Model specification : utility specification

- Leader follower (**constrained**) :

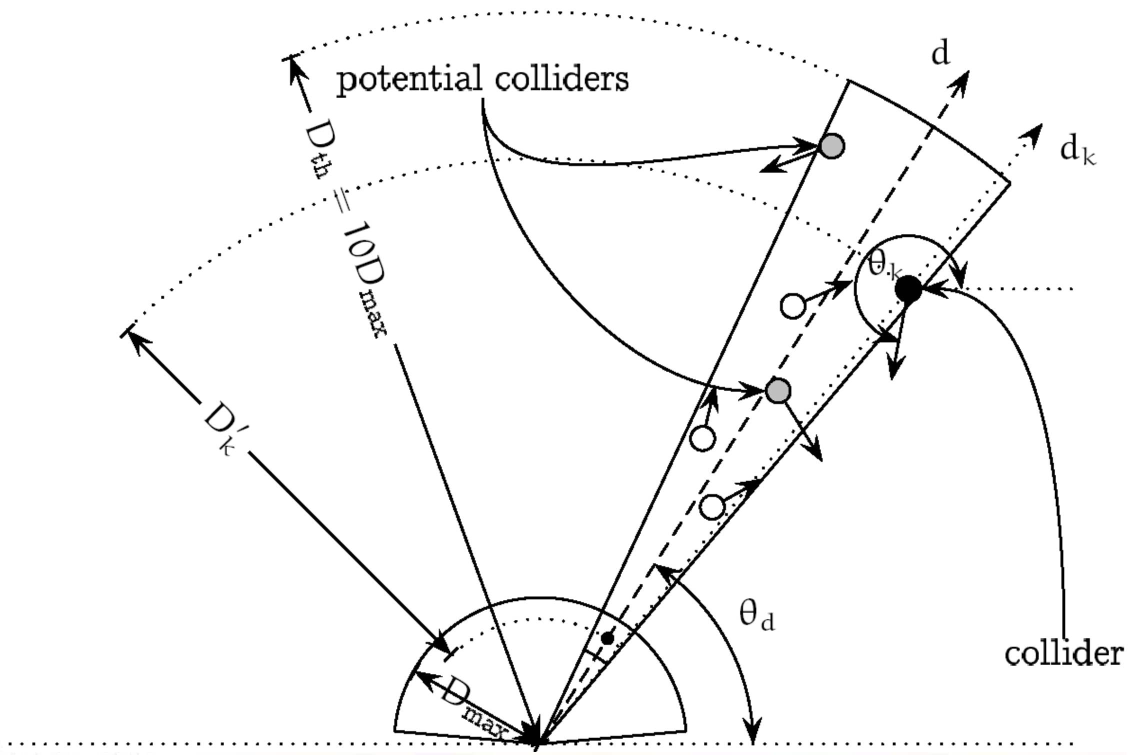
$$I_{v,acc} I^L_{acc} \underbrace{\overset{0.942}{\alpha^L_{acc}} D_L^{\overset{-0.489}{\rho^L_{acc}}}}_{\text{sensitivity}} \underbrace{\Delta v_L^{\overset{0.625}{\gamma^L_{acc}}} \Delta\theta_L^{\overset{-0.171}{\delta^L_{acc}}}}_{\text{stimulus}} + I_{v,dec} I^L_{dec} \underbrace{\overset{3.69}{\alpha^L_{dec}} D_L^{\overset{-0.663}{\rho^L_{dec}}}}_{\text{sensitivity}} \underbrace{\Delta v_L^{\overset{0.652}{\gamma^L_{dec}}}}_{\text{stimulus}}$$



Model specification : utility specification

-0.00639 0.239 non significative

- Collision avoidance (**constrained**) : $I_{\mathbf{d},d_n} I_C \underbrace{\alpha_C e^{-\rho_C D_C}}_{\text{sensitivity}} \underbrace{\Delta v_C^{\gamma_C} \Delta\theta_C^{\delta_C}}_{\text{stimulus}}$



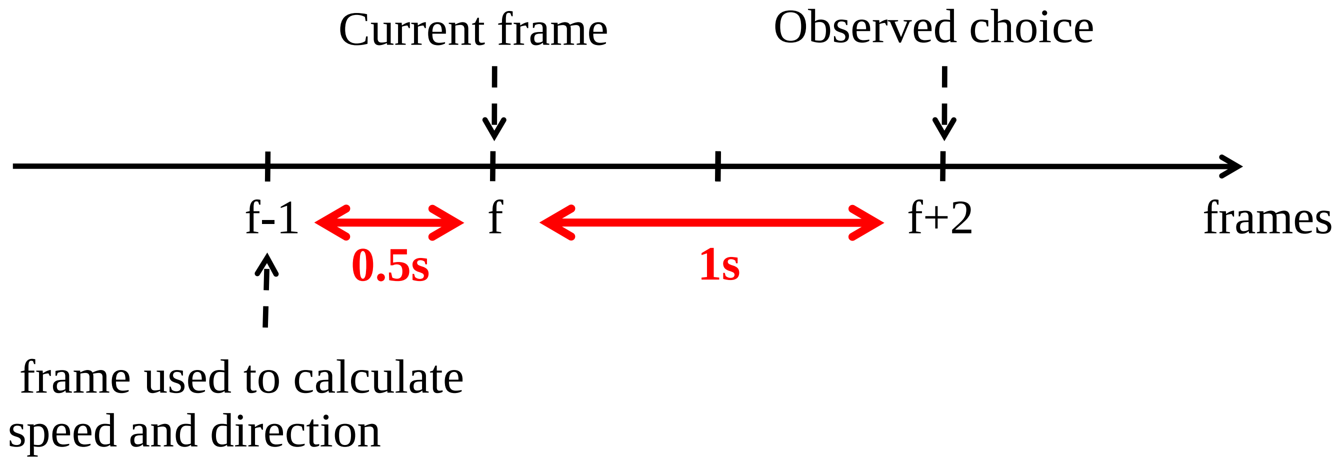
- Introduction
- Discrete choice models
- Model specification
- **Model estimation**
- Model validation
- Simulator
- Conclusion

The Japanese data set : video sequence

- Collected in Sendaï, Japan, on August 2000, large pedestrian **crossing road**

The Japanese data set : data processing

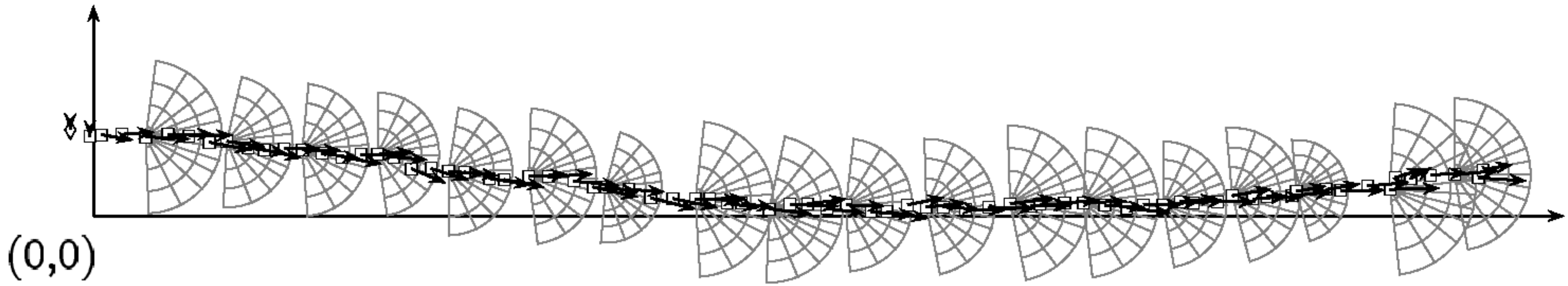
- Tracking from video sequence: **2 observations per second**
- Pedestrians trajectories extracted using 3D-calibration (DLT algorithm)
- For each pedestrian trajectory :



190 pedestrians, 9281 observations

The Japanese data set : pedestrian trajectory

- 4 alternatives are never chosen: 1, 12, 23, 33



Model estimation : general diagnosis

- Estimation made using the free Biogeme package (biogeme.epfl.ch)
- Estimation results :

Number of estimated parameters : 24
Init log-likelihood : -32451
Final log-likelihood : -13997.27
Likelihood ratio test : 36907
$\bar{\rho}^2 = 0.568$

- Parameters values consistent with hypothesis

Model estimation : parameters values

Variable name	Coefficient estimate	t test 0	t test 1
β_{ddir}	-0.0790	-24.53	
β_{ddist}	-1.55	-11.66	
$\beta_{\mathrm{dir\_extreme}}$	-0.0326	-9.30	
$\beta_{\mathrm{dir\_side}}$	-0.0521	-21.87	
$\beta_{\mathrm{dir\_central}}$	-0.0252	-8.74	
β_{accLS}	-4.97	-22.61	
β_{accHS}	-7.47	-5.21	
β_{dec}	-0.0630	-2.40	
λ_{accLS}	4.16	15.94	
λ_{accHS}	0.358	2.09	
λ_{dec}	-2.41	-8.43	
$\alpha^{\mathrm{L}}_{\mathrm{acc}}$	0.942	2.28	
$\rho^{\mathrm{L}}_{\mathrm{acc}}$	-0.489	-2.19	
$\gamma^{\mathrm{L}}_{\mathrm{acc}}$	0.625	2.87	
$\alpha^{\mathrm{L}}_{\mathrm{dec}}$	3.69	6.90	
$\rho^{\mathrm{L}}_{\mathrm{dec}}$	-0.663	-7.11	
$\gamma^{\mathrm{L}}_{\mathrm{dec}}$	0.652	6.19	
$\delta^{\mathrm{L}}_{\mathrm{acc}}$	-0.171	-2.33	
α_C	-0.00639	-9.82	
ρ_C	0.239	-8.28	
μ_{acc}	1.66	9.73	3.88
μ_{const}	1.50	13.46	4.48
$\mu_{central}$	2.35	1.93	1.11
$\mu_{not\_central}$	1.75	9.46	4.04

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

- Introduction
- Discrete choice models
- Model specification
- Model estimation
- **Model validation**
- Simulator
- Conclusion

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model validation : methodology

- Validation of the specification :
  - Developpment of a model with constants only (ASC model)
  - Simulation on the Japanese data set
  - Cross validation on the Japanese data set
- Validation of the model :
  - Simulation on an experimental Dutch data set, **not used for model estimation**
  - Comparison of the proposed model with the ASC model

Model validation : model constants-only

- The simplest model : utility of each alternative represented only by an alternative specific constant (ASC)

- This model with only constants (ASC model) estimated on the Japanese data set.

  ⇒ **28 parameters (33, minus 4 never chosen, minus 1 for normalization)**

- It reproduces the aggregated observations proportions of the Japanese data set

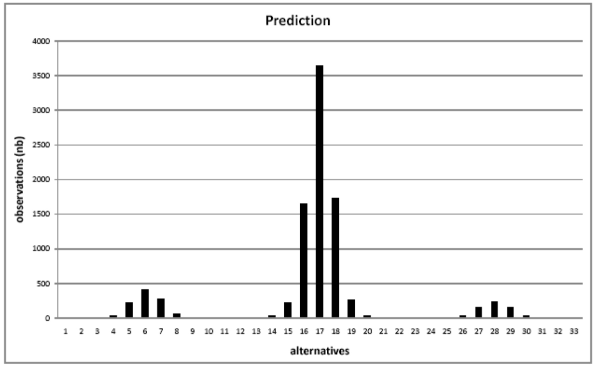
- The ASC model **used for comparison** (for example the number of outliers)

TRANSP-OR

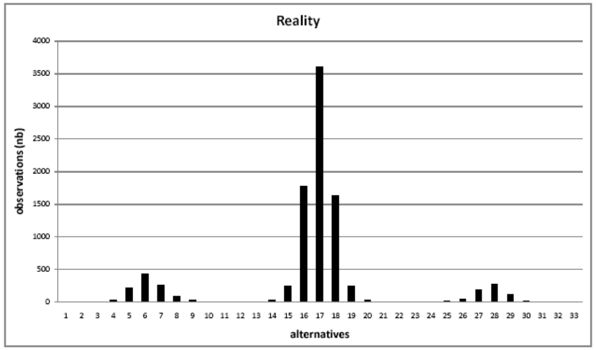
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model validation : simulation on the Japanese data set (Aggregate level)

- The proposed model is applied to the Japanese data set (used for estimation)



(a) Predicted shares



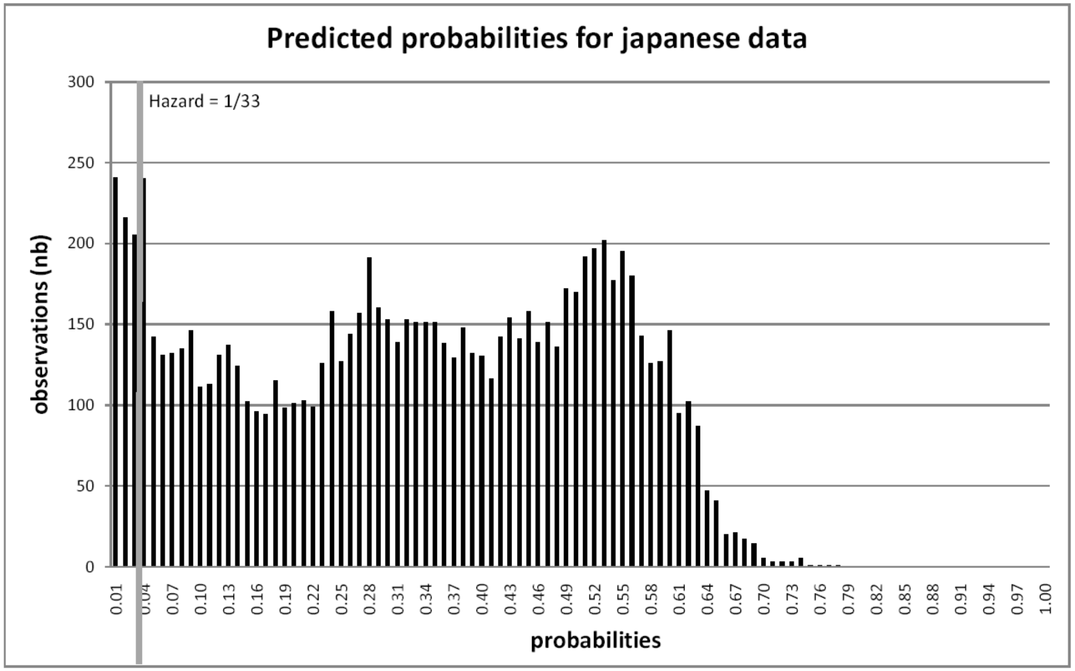
(b) Observed shares

Cone	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
Front	$5-7, 16-18, 27-29$	8489.27	8481	0.10%
Left	$3, 4, 14, 15, 25, 26$	349.67	367	−4.72%
Right	$8, 9, 19, 20, 30, 31$	415.41.	407	2.08%
Extreme left	$1, 2, 12, 13, 23, 24$	12.29	10	22.96%
Extreme right	$10, 11, 21, 22, 32, 33$	14.30	16	−10.59%

Area	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
acceleration	$1-11$	1041.50	1065	−2.21%
constant speed	$12-22$	7606.49	7565	0.55%
deceleration	$23-33$	633.02	651	−2.76%

Model validation : simulation on the Japanese data set (Disaggregate level)

- **Outlier** : Observation with predicted probability less than 1/33 (hazard)



Number of outliers: **7.13**% for proposed model
19.90% for ASC model

Model validation : Cross-validation on the Japanese data set

- Japanese data splited into 5 subsets, each containing 20% of the observations

  ⇒ 5 experiments : 1 subset saved for **validation**; **estimation** of the model on the 4 remaining

- Number of **outliers** (compared with the ASC model cross validation)

Model	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5
Proposed spec.	8.78%	6.36%	7.60%	7.87%	5.87%
Constant only	20.79%	20.70%	17.13%	19.88%	18.64%

⇒ **Robust specification**

TRANSP-OR

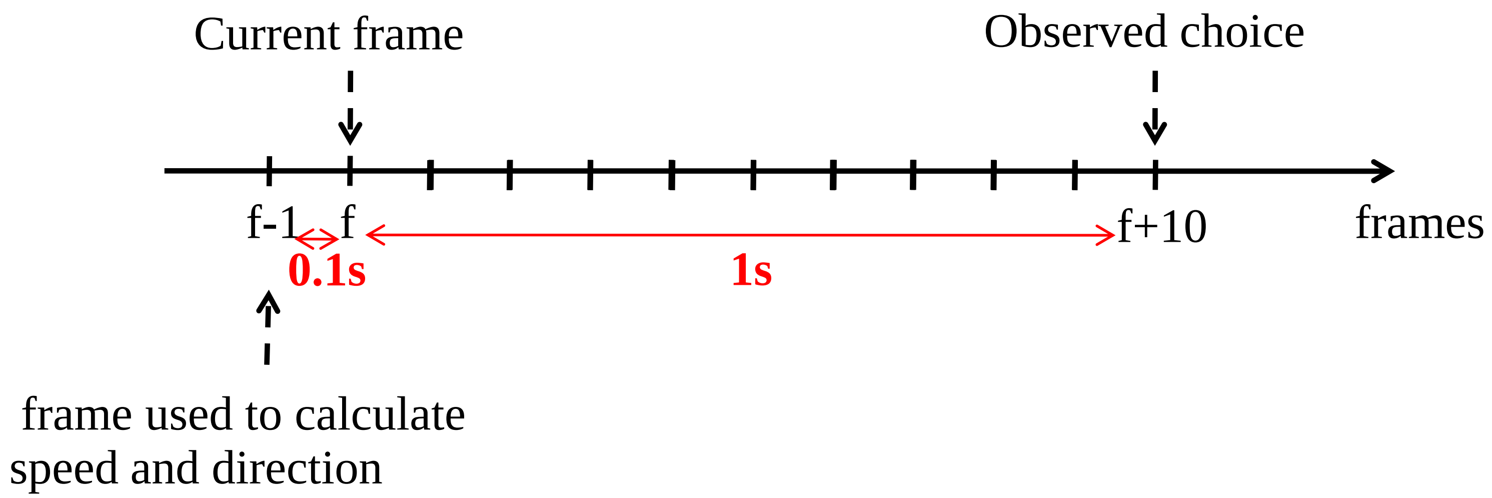
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

The Dutch data set : video sequence

- Collected at Delft university, in 2000-2001, 2 pedestrians crossing flows

The Dutch data set : general information

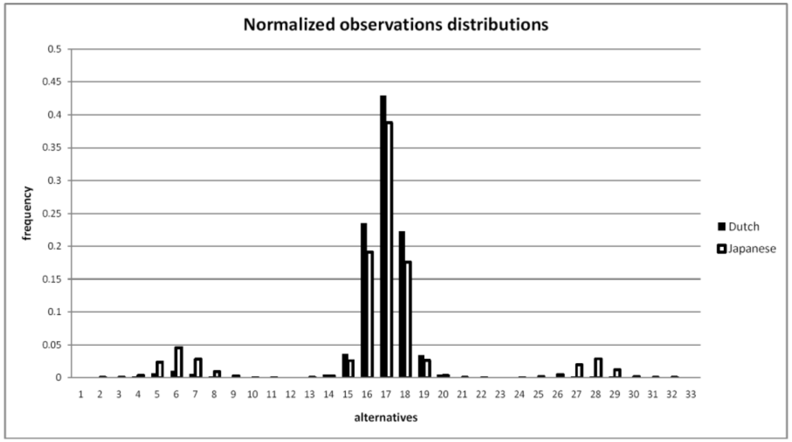
- **Experimental** data set
- Video sequence recorded at **10 frames per second**
- Pedestrians trajectories extracted from the video sequence
- For each pedestrian trajectory :



724 pedestrians, 47481 observations

The Dutch data set : comparison with the Japanese data set

- Normalized observations distribution among alternatives



- Observations repartitions inside the nest (Japanese / Dutch)

Nest	# steps	% of total
acceleration	1065	11.48%
constant speed	7565	81.51%
deceleration	651	7.01%
central	4297	46.30%
not central	4984	53.70%

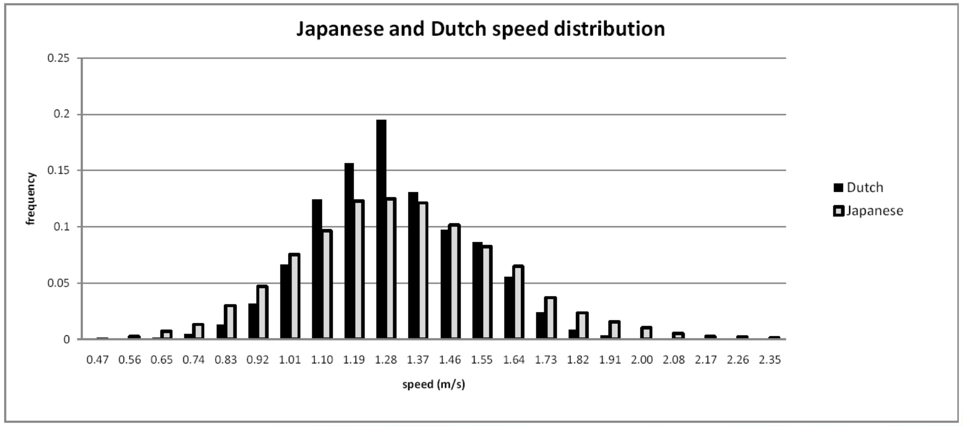
Nest	# steps	% of total
acceleration	1273	2.68%
constant speed	45869	96.61%
deceleration	339	0.71%
central	20950	44.12%
not central	26531	55.88%

The Dutch data set : comparison with the Japanese data set

- Quite similar observations proportions in the **direction's cones** (not for speed regime)

Dataset	extremeleft	left	front	right	extremeright
Japanese	0.11%	3.95%	91.38%	4.39%	0.17%
Dutch	0.06%	4.40%	91.35%	4.15%	0.04%

- Speed distributions have different shapes (experimental design of Dutch data set)

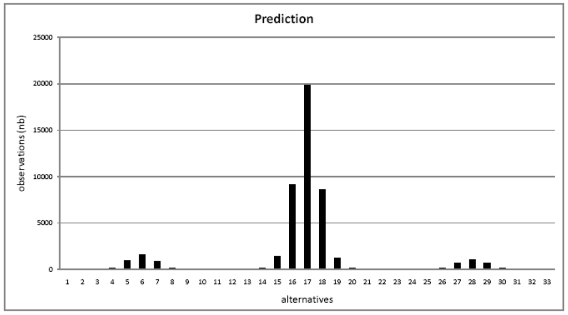


TRANSP-OR

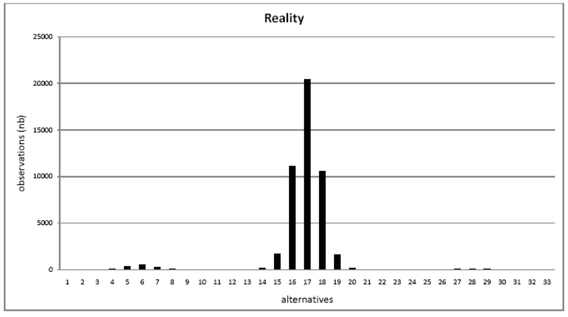
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model validation : simulation on the Dutch data set (Aggregate level)

- The proposed model is applied to the **Dutch** data set (**NOT** used for estimation)



(a) Predicted



(b) Observed

Cone	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
Front	$5-7, 16-18, 27-29$	43619.98	43374	0.57%
Left	$3, 4, 14, 15, 25, 26$	1968.79	2089	−5.75%
Right	$8, 9, 19, 20, 30, 31$	1764.39	1972	−10.53%
Extreme left	$1, 2, 12, 13, 23, 24$	45.86	27	69.85%
Extreme right	$10, 11, 21, 22, 32, 33$	81.97	19	331.44%

Area	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
acceleration	$1-11$	3892.35	1273	205.76%
constant speed	$12-22$	40733.53	45869	−11.20%
deceleration	$23-33$	2855.12	339	742.22%

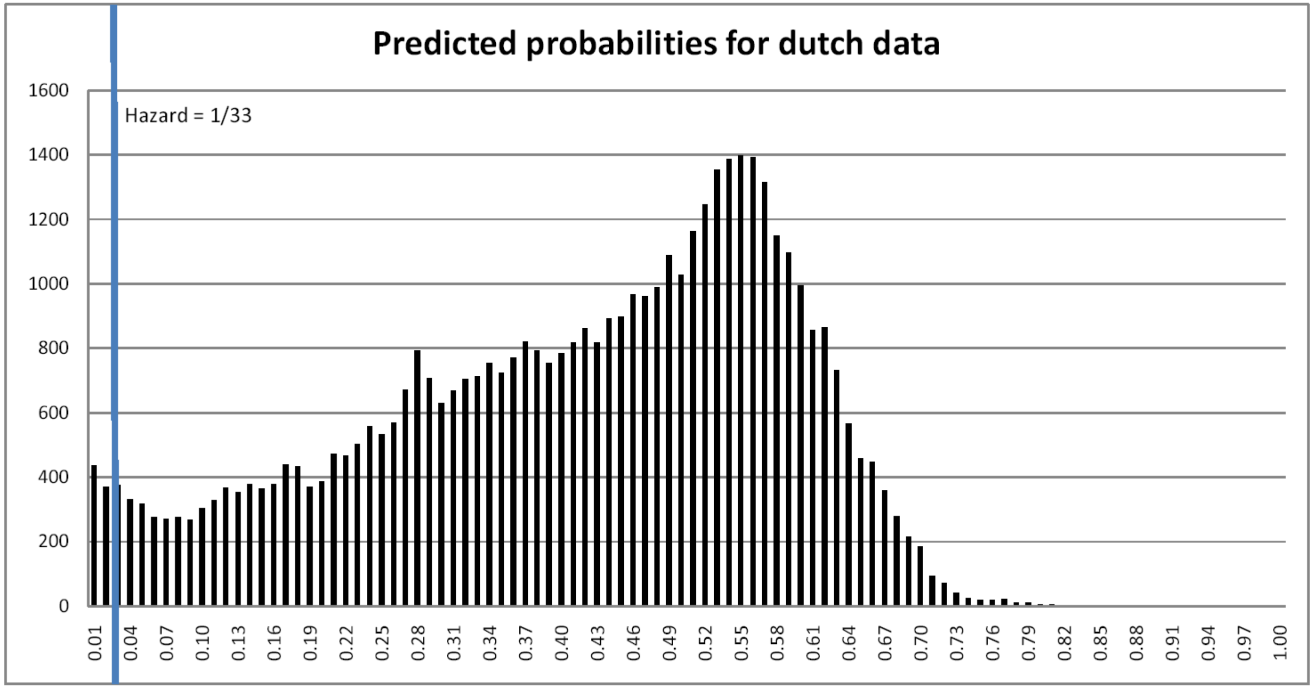
Overprediction of acceleration and deceleration

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model validation : simulation on the Dutch data set (Disaggregate level)

- **Outlier** : Observation with predicted probability less than 1/33 (hazard)



Number of outliers: **2.48**%

Model validation : Comparison with the ASC model on the Dutch data set (Aggregate level)

- The ASC model is applied to the Dutch data set and compared to the proposed model)

ASC model

Cone	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
Front	$5-7, 16-18, 27-29$	43386.42	43374	0.03%
Left	$3, 4, 14, 15, 25, 26$	1877.47	2089	−10.13%
Right	$8, 9, 19, 20, 30, 31$	2082.10	1972	5.58%
Extreme left	$1, 2, 12, 13, 23, 24$	51.16	27	89.47%
Extreme right	$10, 11, 21, 22, 32, 33$	81.85	19	330.80%

Area	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
acceleration	$1-11$	5448.24	1273	327.98%
constant speed	$12-22$	38700.42	45869	−15.63%
deceleration	$23-33$	3330.34	339	882.40%

Proposed model

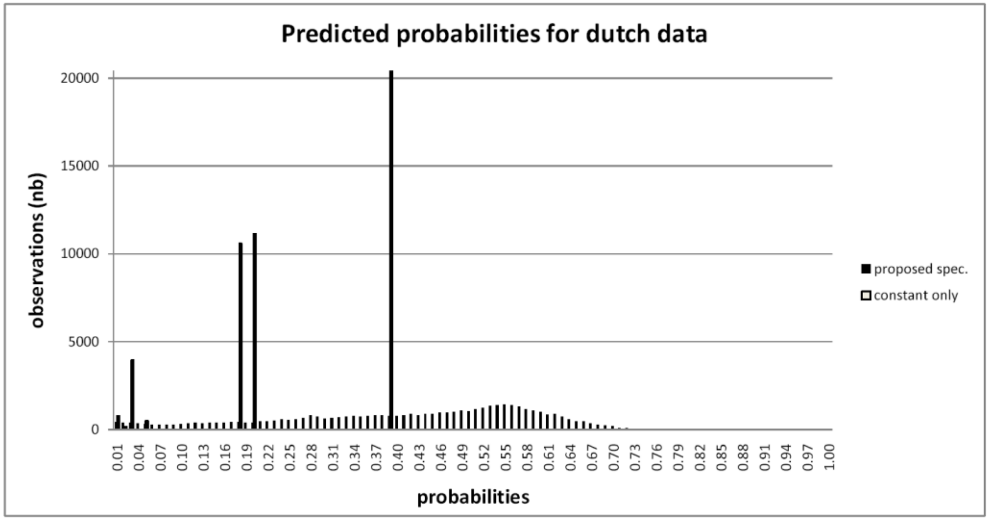
Cone	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
Front	$5-7, 16-18, 27-29$	43619.98	43374	0.57%
Left	$3, 4, 14, 15, 25, 26$	1968.79	2089	−5.75%
Right	$8, 9, 19, 20, 30, 31$	1764.39	1972	−10.53%
Extreme left	$1, 2, 12, 13, 23, 24$	45.86	27	69.85%
Extreme right	$10, 11, 21, 22, 32, 33$	81.97	19	331.44%

Area	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
acceleration	$1-11$	3892.35	1273	205.76%
constant speed	$12-22$	40733.53	45869	−11.20%
deceleration	$23-33$	2855.12	339	742.22%

→ **Equivalent for direction** (logical, due to proportions)

Model validation : simulation on the Japanese data set (Disaggregate level)

- **Outlier** : Observation with predicted probability less than 1/33 (hazard)



Number of outliers:
- **2.48**% for proposed model
- **10.31**% for ASC model

Superiority of the proposed model

- Introduction
- Discrete choice models
- Model specification
- Model estimation
- Model validation
- **Simulator**
- Conclusion

TRANSP-OR

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Simulator

- **Implementation** of the **developped specification** in a simulator
- Simulation of **2 pedestrian crossing flows** with the model
- Examples : | - Simulation of 300s
  - Start : random speed and direction
  - Finish : random destination

Ex1 : low density, 2 pedestrians per second entering

Ex2 : high density, 5 pedestrians per second entering

TRANSP-OR

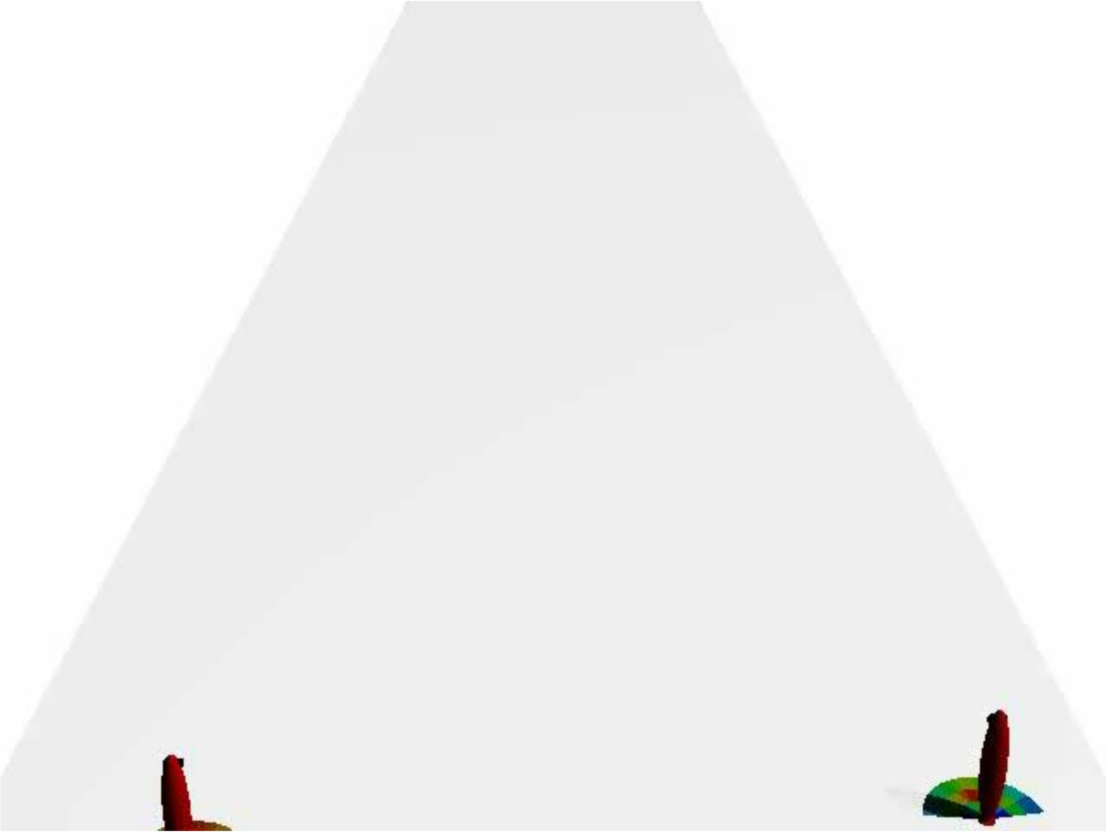
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Simulator

- **Low** density :

Simulator

- **High** density :



Conclusions and Perspectives

- **Conclusions :**
  - Discrete choice model for pedestrian walking behavior with ‘**unconstrained**‘ and ‘**constrained**’ parameters
  - Model **estimated** on a real data set, parameters values consistent with hypothesis
  - Model validated on a real data set, **not used for estimation**
  - Operating **Simulator**
- **Perspectives :**
  - Improve the **acceleration** and **deceleration** patterns
  - Incorporate **physical characteristics** of the pedestrians
  - Model the **strategical** and **tactical** behavioural levels

Thanks for your attention

http://transp-or2.epfl.ch/publications.php#techrep