

Hybrid Modeling of a Bike-Sharing Transportation System

J. Raimbault^{1,2}

¹ UMR 8504 CNRS - Géographie-Cités, ² UMR-T 9403 IFSTTAR - LVMT



Introduction

- ▶ Bike-sharing systems are presented as a new sustainable urban transportation mode [Midgley, 2009].
- ▶ Already well-studied from a *top-down* point of view, e.g. to unveil urban mobility patterns through datamining [O'Brien et al., 2013, Borgnat et al., 2009], or to optimize system design by Operational Research [Lin and Yang, 2011].
- ▶ We propose a hybrid *bottom-up* approach, using statistical analysis to parametrize an agent-based model of a bike-sharing system dynamics in the spirit of [Kohler and Reese, 2014]. It aims to exploit “poor” data sources (methodological contribution) and to test the role of user-centered parameters (thematic contribution).

Model

Simple agent-based model proposed in [Raimbault, 2015b].

Agents : bikers with information $i(\mathbf{b})$ (boolean), tolerated walking radius $r(\mathbf{b})$ and mean speed $\bar{v}(\mathbf{b})$; docking stations located in space with current standing bikes $p_b(\mathbf{s}, t)$ and capacity $c(\mathbf{s})$.

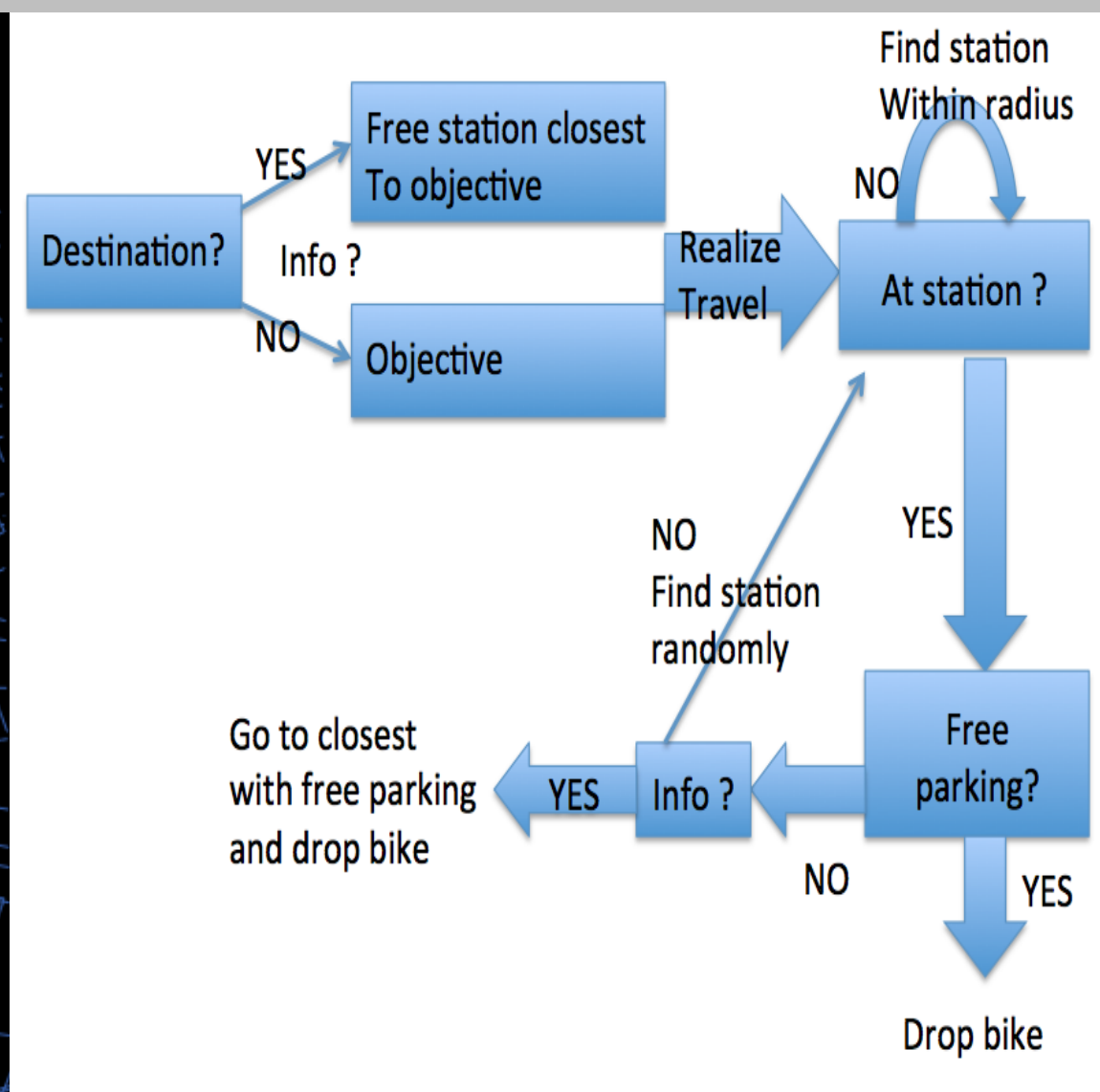
Environment : Road network, of which stations are on nodes and where movement of bikers is embedded ; temporal fields of origin $O(\mathbf{t})$ and destination $D(\mathbf{t})$ (probabilities of O/D given a trip), boundaries conditions $N(\mathbf{t})$ as flows (in- and outflows) at fixed boundaries points, are given by parametrization.

At each time step:

1. Start new travels randomly using O, D, N
2. Make bikers in travel advance of the corresponding distance
3. Finish travels and redirect bikers when needed (see flowchart of bikers behavior)

Indicators : *Model fit* : Mean-square error on load factors ; *System performance* : Spatial heterogeneity of bike distribution, number of adverse events, quantity of detours by bikers.

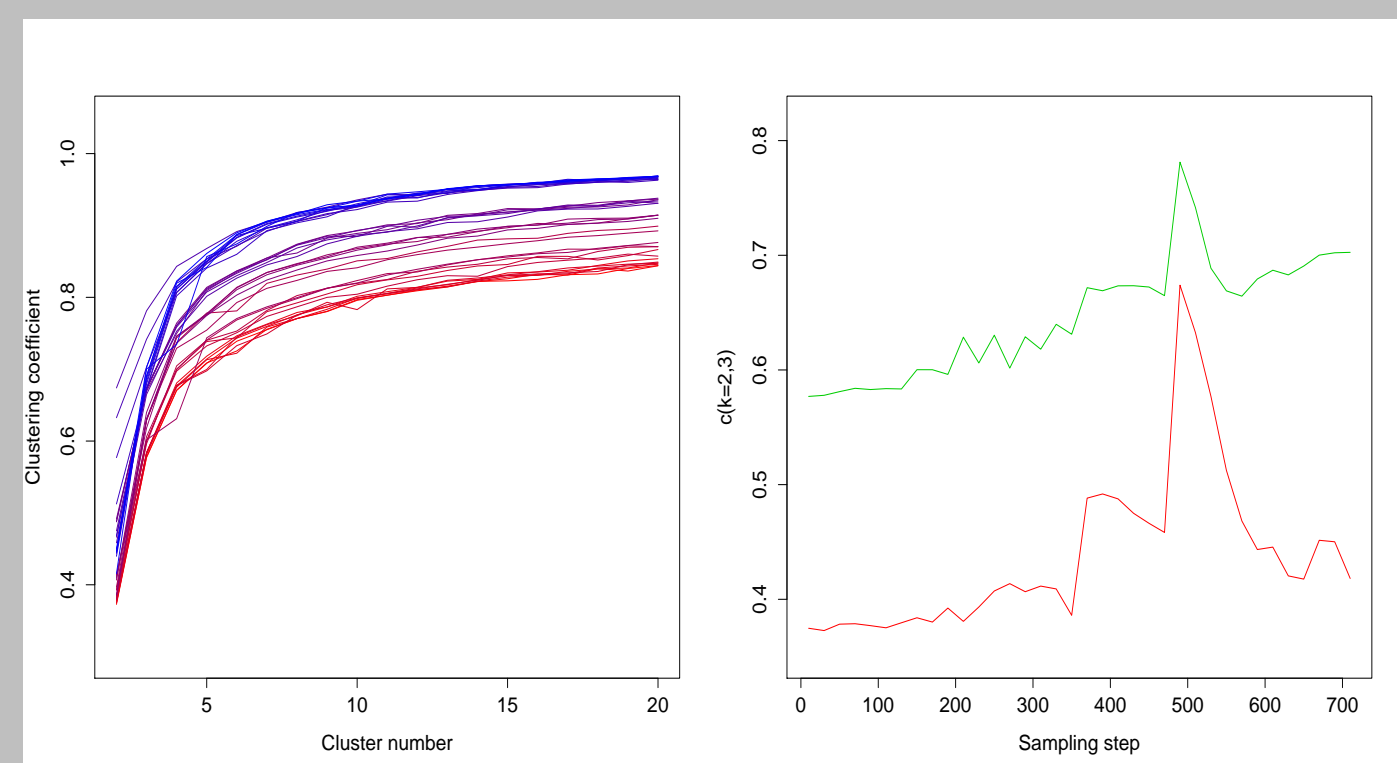
Implemented in NetLogo coupled with R for statistical analysis [Raimbault, 2015a].



Left : Example of graphical output of the model. Right : Flowchart of bikers behavior.

Statistical Analysis of Raw Data for Parametrization

Raw Data on Paris' bike-sharing system dynamics provided publicly by operator, collected each 5s during 6 month. “Poor” data : only docking stations status, allows to extract arrivals and departures.



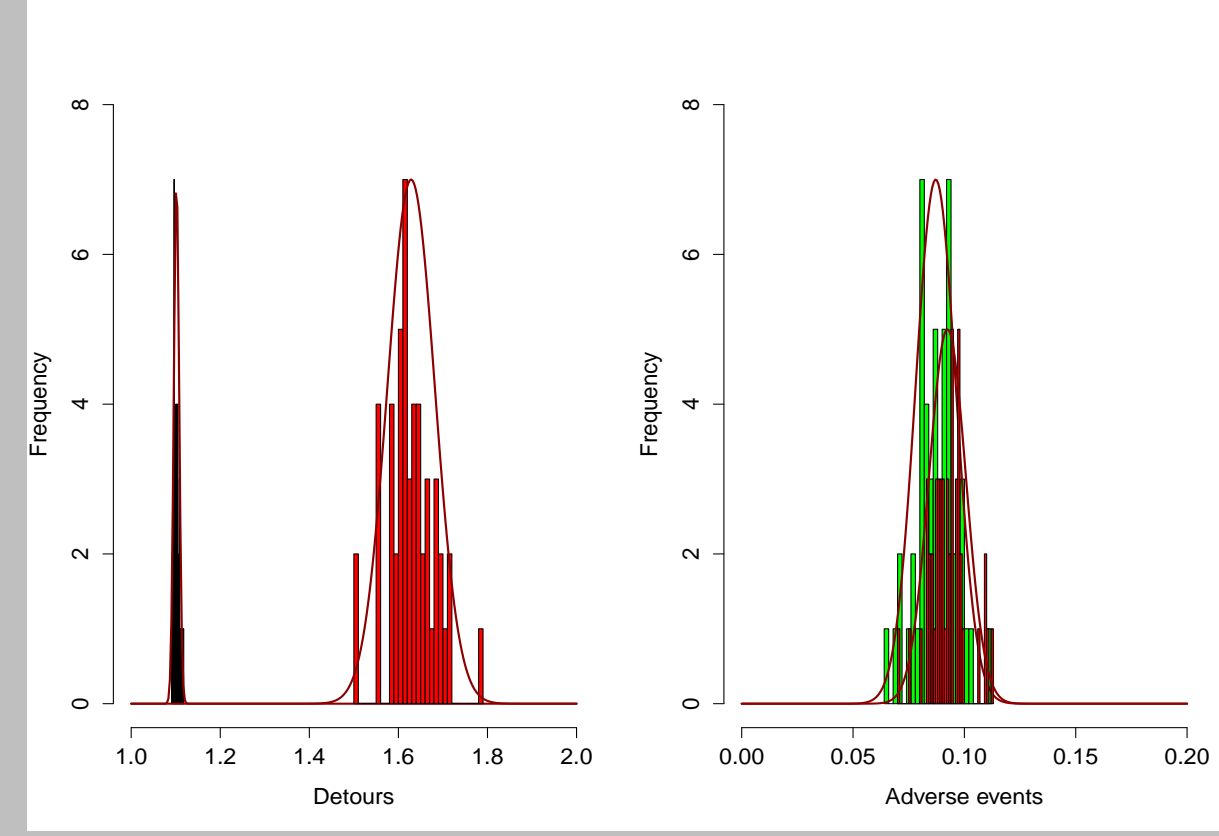
Clustering Analysis of raw data to isolate the patterns of a typical day. Dimension reduction by clustering on stations (Left : role of cluster number) and time-series sampling (Right : effect of sampling step on information).

O/D fields are then estimated by a multi-kernel estimation, similarly to Geographically Weighted Regression [Brunsdon et al., 2002], with $(d_i(t))$ real arrivals at $(\vec{x}_i(t))$,

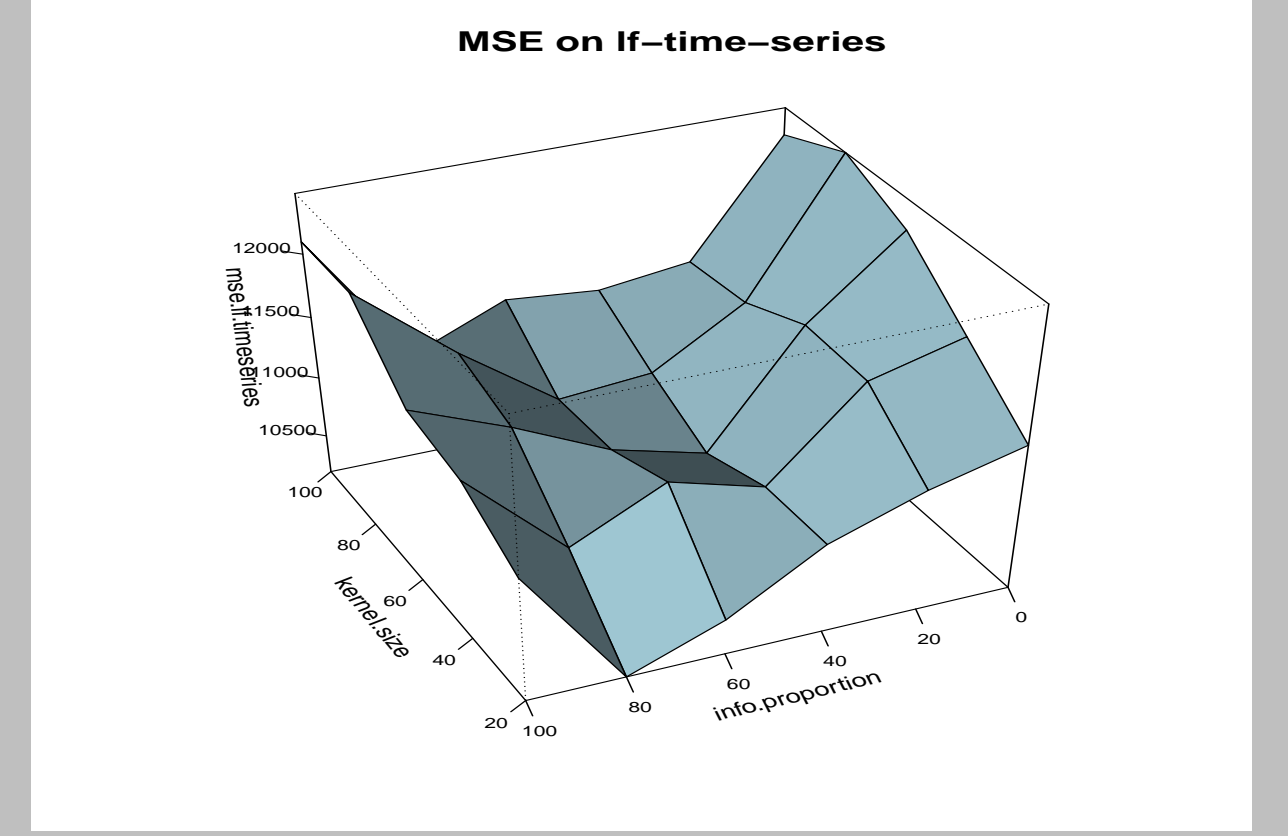
$$[D(t)](\vec{x}) = \frac{1}{K} \sum_i d_i(t) \cdot \exp\left(\frac{\|\vec{x} - \vec{x}_i\|}{2\sigma^2}\right)$$

Validation and Calibration

Reduced parameter space to be explored contains p_I proportion of user having access to information, \bar{r} mean tolerated walking radius and σ kernel size. Grid exploration with repetitions, use of OpenMole software for systematic model exploration [Reuillon et al., 2013].



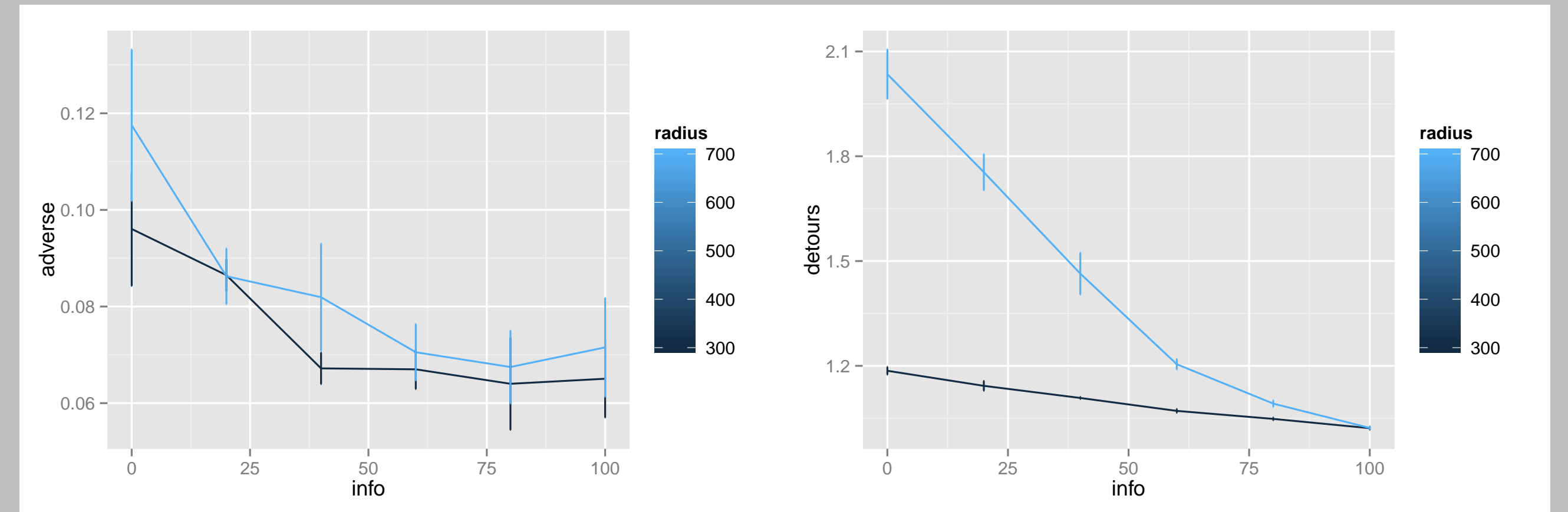
Internal validation : distributions of indicators for different parameter values.



Calibration : minimization of MSE on (p_I, σ) plane.

Application : User-based Improvement Strategy

We test the influence of user parameters such as quantity of information or propensity to walk on the performances of the system.



Influence of information proportion on adverse events (left) and on quantity of detours (right) for different values of walking radius. We observe a saturation of performance improvement at a low value $p_I \simeq 30\%$.

Extension : Discrete Choices Model for User Behavior

- ▶ Flexibility of model allows to test other rules for user behavior. We extend the model with a Discrete Choices module for the choice between waiting (w) and moving (m) on at a full docking station, where utilities are, with t_w expected waiting time, $d(i)$ expected detour, $\tilde{d}(i)$ expected distance difference to destination,

$$U_w(i) = \beta_t t_w + \beta_d \tilde{d}(i) \cdot i + \varepsilon_w \text{ and } U_m(i) = \beta_t \frac{d(i)}{\bar{v}} + \beta_d \tilde{d}(i) \cdot i + \varepsilon_m$$

- ▶ Data from a questionnaire [Bourcet et al., 2014] aimed to estimate above DC parameters (among others) but with poorly significant estimates ($N = 150$), to obtain their feasible space.
- ▶ Extended parameter space can then be explored, calibration allowing to extract better indirect estimates of DC parameters. First results of directed gradient and grid explorations give rather “chaotic” MSE surfaces, would need finer grain exploration (computational limit).

Conclusion

- ▶ Simple hybrid model gives interesting thematic results, confirming the potential of “poor” datasets.
- ▶ Methodology of indirectly estimating DC parameters through ABM calibration could be generalized to other types of system.

References

- Borgnat, P., Abry, P., and Flandrin, P. (2009). Modélisation statistique cyclique des locations de vélo à lyon. In *XXII colloque GRETSI (traitement du signal et des images)*, Dijon (FRA), 8-11 septembre 2009. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images.
- Bourcet, M., Lesturgie, J., Allain, T., Etienne, C., Sebes, A., and Raimbault, J. (June 2014). Le vélo comme choix de mode. Technical report, ENPC, D'épartement VET.
- Brunsdon, C., Fotheringham, A., and Charlton, M. (2002). Geographically weighted summary statistics—a framework for localised exploratory data analysis. *Computers, Environment and Urban Systems*, 26(6):501–524.
- Kohler, T. A. and Reese, K. M. (2014). Long and spatially variable neolithic demographic transition in the north american southwest. *Proceedings of the National Academy of Sciences*, page 201404367.
- Lin, J.-R. and Yang, T.-H. (2011). Strategic design of public bicycle sharing systems with service level constraints. *Transportation research part E: logistics and transportation review*, 47(2):284–294.
- Midgley, P. (2009). The role of smart bike-sharing systems in urban mobility. *JOURNEYS*, 2:23–31.
- O'Brien, O., Cheshire, J., and Batty, M. (2013). Mining bicycle sharing data for generating insights into sustainable transport systems. *Journal of Transport Geography*.
- Raimbault, J. (2015a). Bike sharing abm implementation at <https://github.com/justeraimbault/discretechoicesbikesharing/tree/master/models/abm>.
- Raimbault, J. (2015b). User-based solutions for increasing level of service in bike-sharing transportation systems. In *Complex Systems Design & Management*, pages 31–44. Springer.
- Reuillon, R., Leclaire, M., and Rey-Coyrehourcq, S. (2013). Openmole, a workflow engine specifically tailored for the distributed exploration of simulation models. *Future Generation Computer Systems*, 29(8):1981–1990.