Coupling heterogeneous urban models

J. Raimbault[1,2,3,4,*] and M. Batty[2]
*juste.raimbault@ign.fr

[1]LASTIG, Univ Gustave Eiffel, IGN-ENSG
[2]CASA, UCL
[3]UPS CNRS 3611 ISC-PIF
[4]UMR CNRS 8504 Géographie-cités

Applied Urban Modelling 2022 Symposium
Session 11: Advances in model integration
30/06/2020

Coupling heterogeneous urban models

1. Integration of urban models
2. Modular open transport models
3. Anticipating the impact of future transportation infrastructure
4. Discussion

Modeling urban complexity

Large scale urban models are intrinsically flawed and do not reach their goals of long-term application to planning: **Requiem for large scale models in 1973** [Lee Jr, 1973]

Urban analytics and Smart Cities approaches may follow the same path if they ignore the past and the complexity of cities [Batty, 2014]

To foster relevance of large urban models:

- Transparency on data and implementation, reproducibility
- Validation of models and sub-models: from small simple models well validated to larger integrated models

→ Open, reproducible urban models can be shared, coupled into modular integrated models, tested and validated [Banos, 2013]

Proposed research framework

Sustainable urban systems: (i) multiple contradictory objectives; (ii) implemented by stakeholders at different scales, within various information and power contexts; (iii) adaptive on multiple time scales.

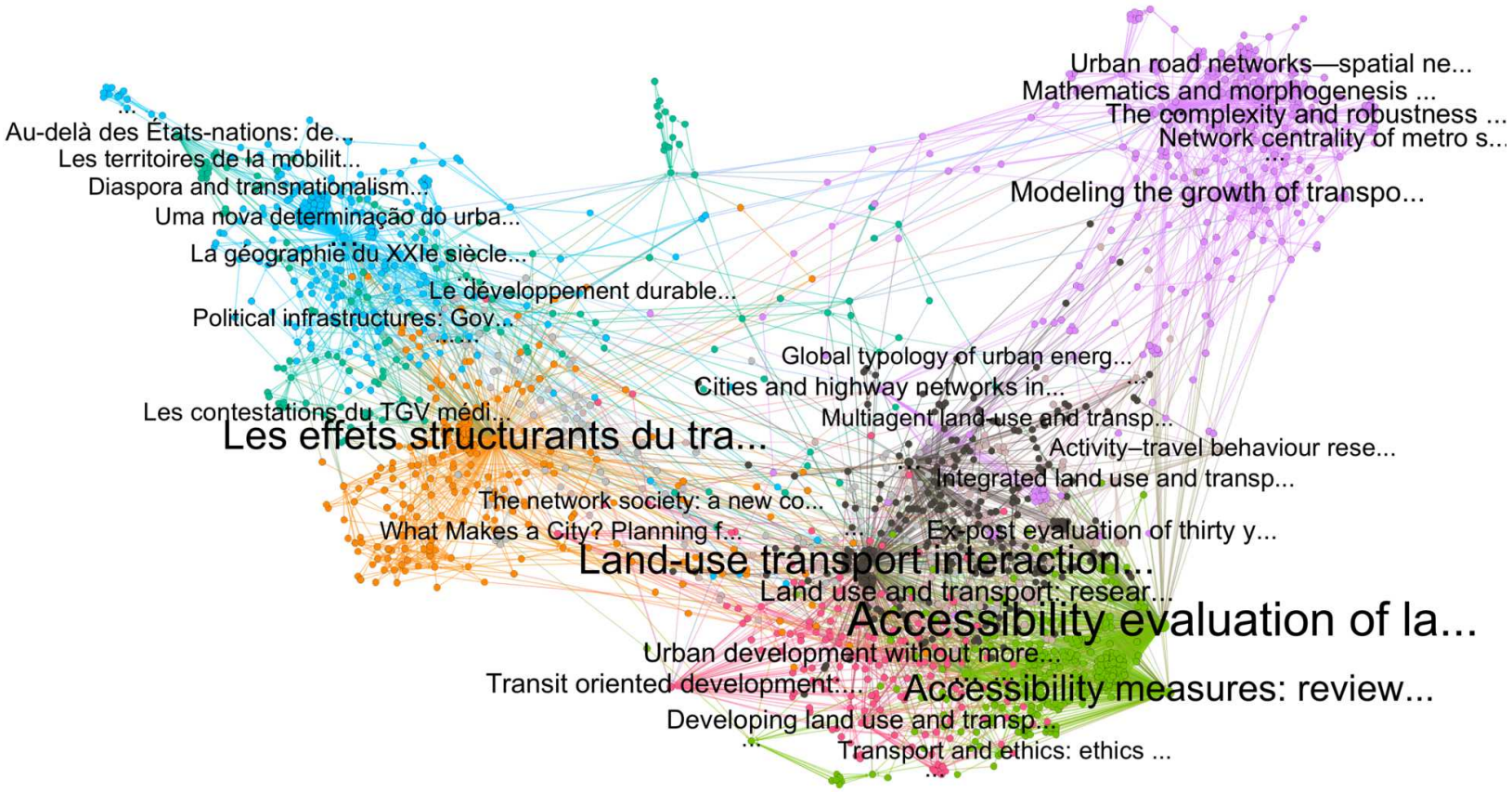
Models are essential tools to (i) capture complexity; (ii) construct integrated perspectives; (iii) link data, empirical stylised facts and decision-making.

→ **Integrated models** to simulate multiple dimensions of **urban systems** towards decision-making in the context of **sustainable transitions**.

- Horizontal integration (model coupling and interdisciplinarity)
- Vertical integration (multi-scale models)
- Model exploration and validation methods

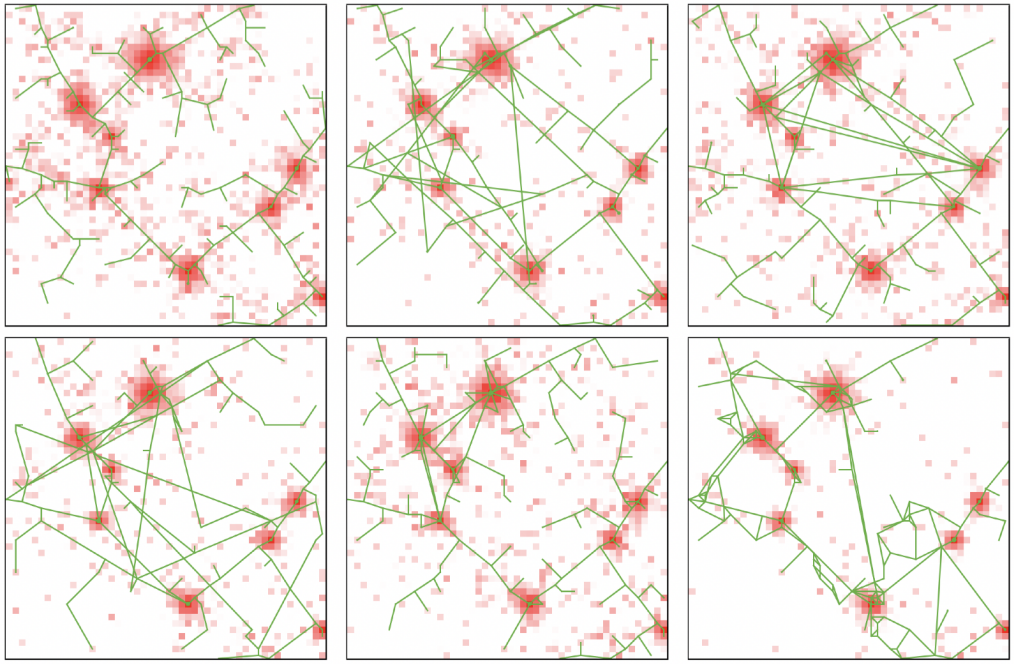
Horizontal integration: interdisciplinarity

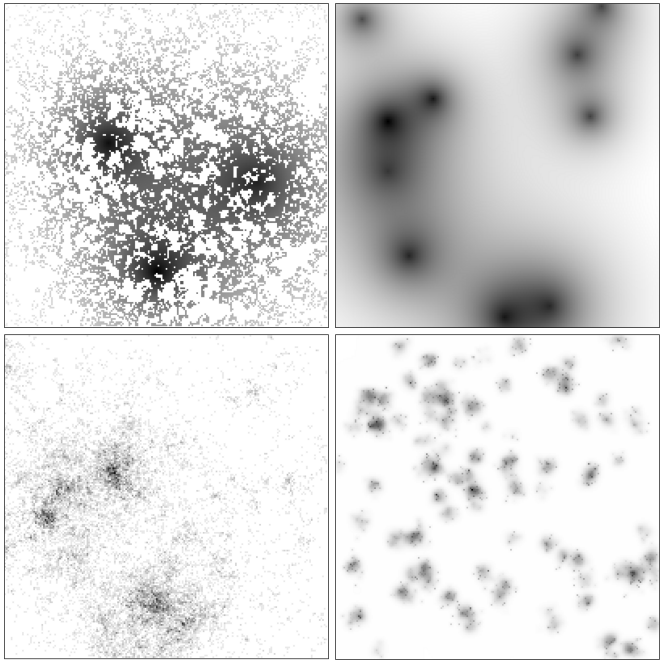
Literature mapping and systematic review tools to enhance integration



Raimbault, J. (2019). Exploration of an interdisciplinary scientific landscape. Scientometrics, 119(2), 617-641.

Horizontal integration: multi-modeling and benchmarks





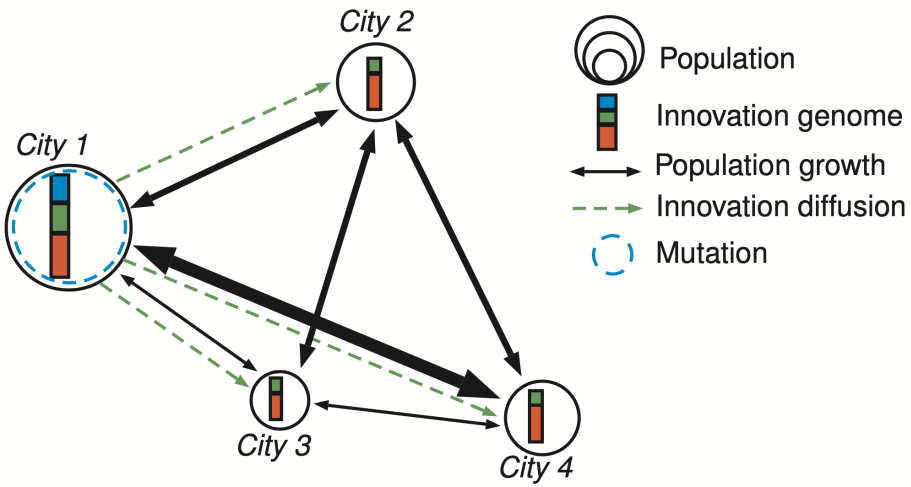
Benchmarking network and urban morphogenesis models

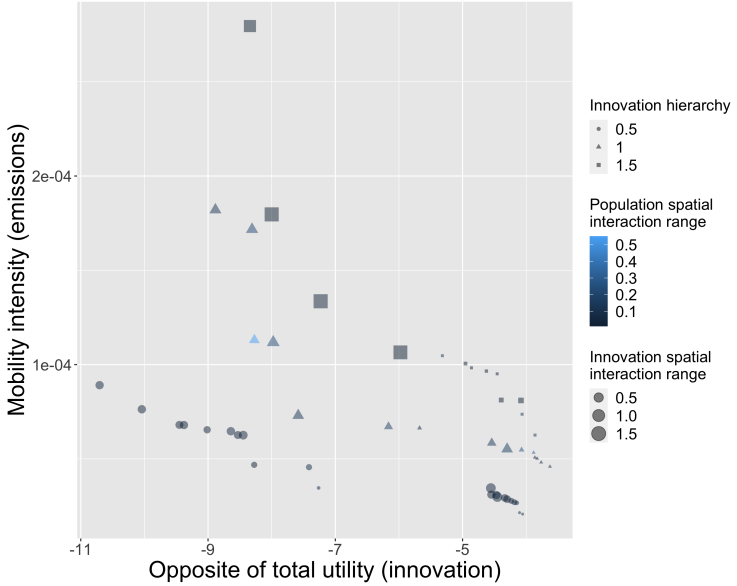
Raimbault, J. (2018). Multi-modeling the morphogenesis of transportation networks. In Artificial Life Conference Proceedings (pp. 382-383). MIT Press, Cambridge.

Raimbault, J. (2020). A comparison of simple models for urban morphogenesis. arXiv preprint arXiv:2008.13277.

Raimbault, J. (2021). Complementarity of generative models for road networks. arXiv preprint arXiv:2109.15206.

Horizontal integration: trade-offs between SDGs



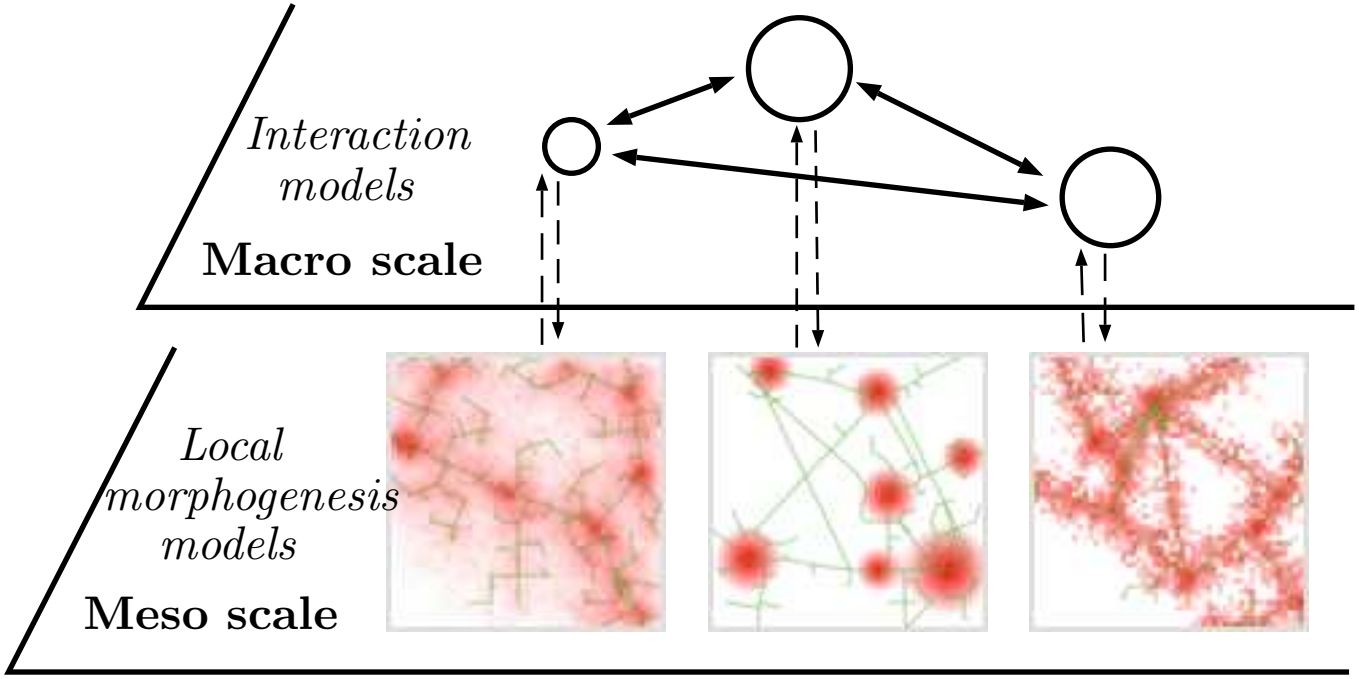


Raimbault, J., & Pumain, D. (2022). Trade-offs between sustainable development goals in systems of cities. Journal of Urban Management.

Work in progress: coupling with an economic exchanges model [Cottineau et al., 2015] and a co-evolution model between transport infrastructure and cities [Raimbault, 2021b]

→ many-objective optimisation of proxy indicators for 5 SDGs (emissions, wealth, economic inequalities, innovation, infrastructure).

Vertical integration: towards multi-scale models



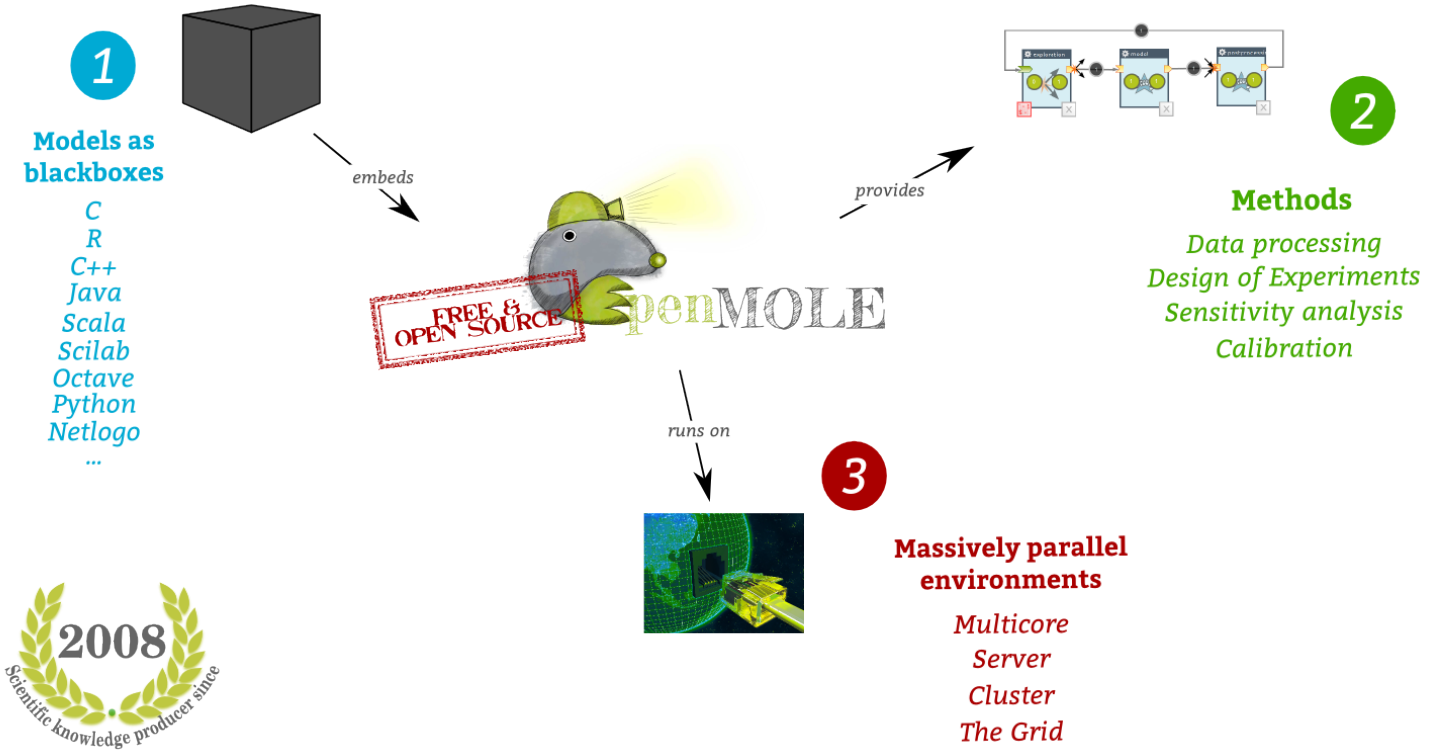
Processes specific to scales, coupling implies dedicated ontologies

Raimbault, J. (2021). Strong coupling between scales in a multi-scalar model of urban dynamics. arXiv preprint arXiv:2101.12725.

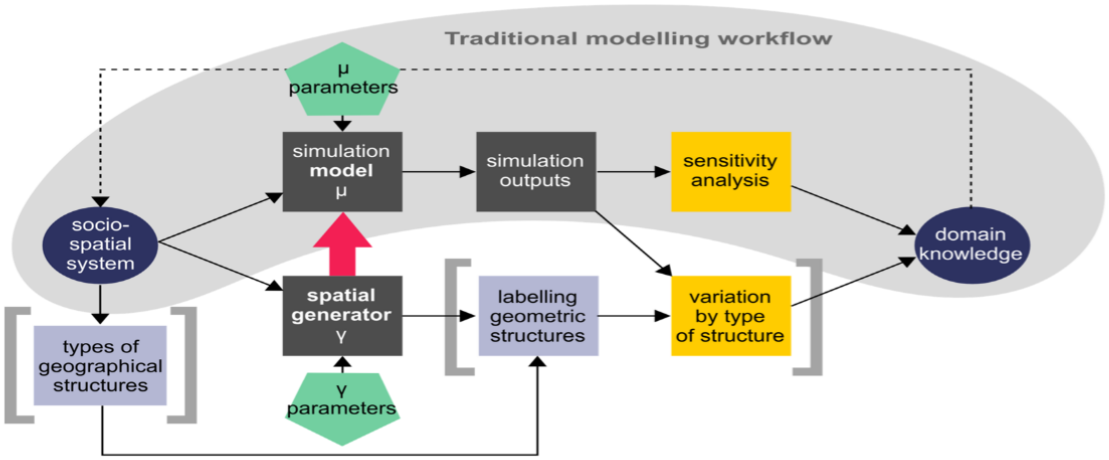
Raimbault, J. (2021). A multiscale model of urban morphogenesis. arXiv preprint arXiv:2103.17241.

Model exploration methods to foster knowledge integration

OpenMOLE software [Reuillon et al., 2013]: *(i) Innovative exploration methods; (ii) Scaling of methods on high performance computing environments; (iii) Scripts to embed and couple models.*

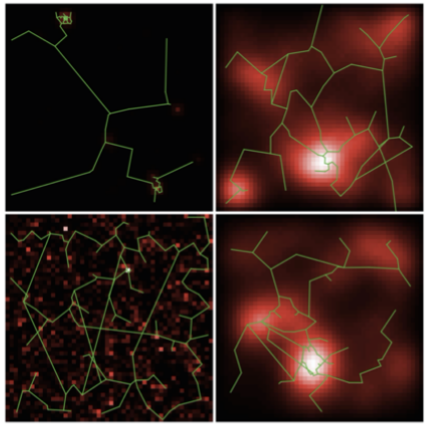


Validation: towards spatial sensitivity analysis

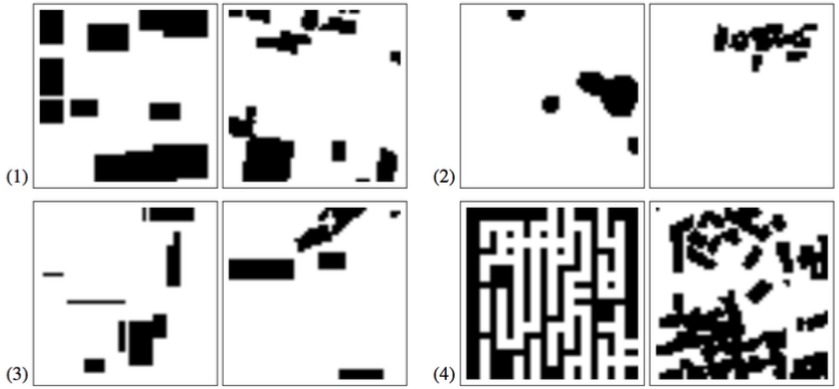


Raimbault, J., Cottineau, C., Le Texier, M., Le Nechet, F., Reuillon, R. (2019). Space Matters: Extending Sensitivity Analysis to Initial Spatial Conditions in Geosimulation Models. *Journal of Artificial Societies and Social Simulation*, *22*(4).

Raimbault, J., Perret, J., & Reuillon, R. (2020). A scala library for spatial sensitivity analysis. *GISRUK 2020 Proceedings, 32.*



Raimbault, J. (2019). Second-order control of complex systems with correlated synthetic data. *Complex Adaptive Systems Modeling*, *7*(1), 1-19.



Raimbault, J., Perret, J. (2019). Generating urban morphologies at large scales. In *Artificial Life Conference Proceedings* (pp. 179-186).

1. Integration of urban models
2. Modular open transport models
3. Anticipating the impact of future transportation infrastructure
4. Discussion

Sensitivity analysis of MATSim

Large scale urban/transport ABMs must be validated for relevant and robust policy applications

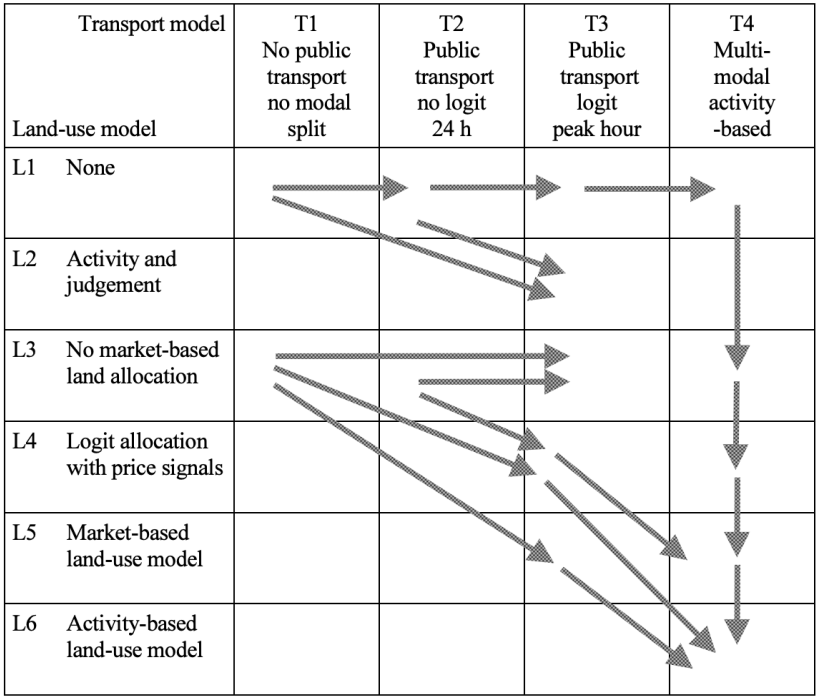
A few examples of MATSim validation or sensitivity analysis in the literature: uncertainty [Bienzeisler et al., 2021], sensitivity analysis [Zhuge et al., 2019], discrete choice parameters [Hörl, 2021]

Research objective:

Provide a modular and open implementation of MATSim generic to any UK urban area and test global sensitivity analysis methods on it

Land-use transport models

Land-use transport models as a progressive complexification through coupling of detailed sub-models

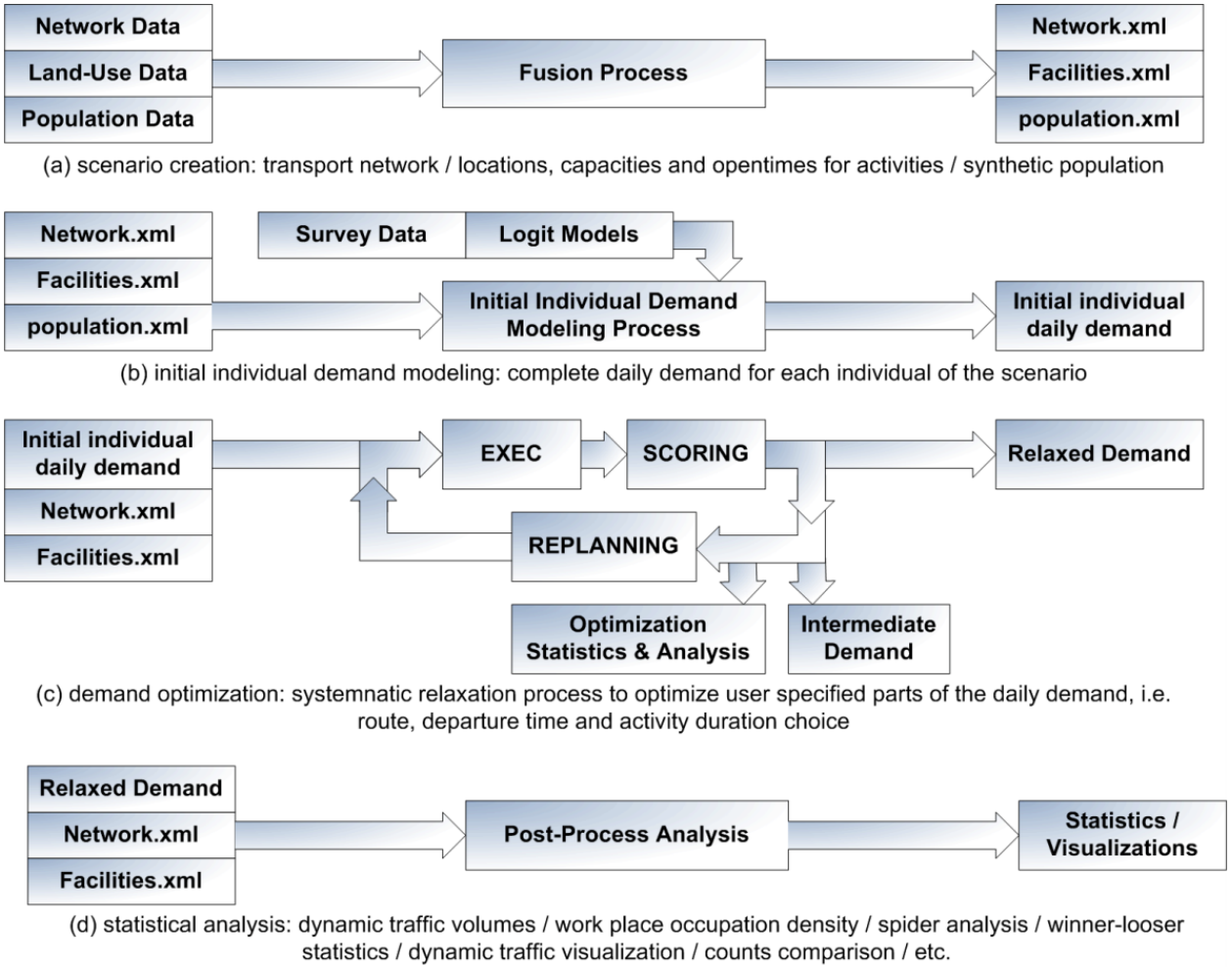


Models	Speed of change							
	Very slow		Slow		Fast		Immediate	
	Networks	Land use	Work-places	Housing	Employ-ment	Popula-tion	Goods transport	Travel
BOYCE	+				+	+		+
CUFM		+	+	+	+	+		
DELTA/START	+	+	+	+	+	+	+	+
HUDS				+	+	+		
IMREL	+	+	+	+	+	+		+
IRPUD	+	+	+	+	+	+		+
ITLUP	+	+			+	+		+
KIM	+				+	+	+	+
LILT	+	+	+	+	+	+		+
MEPLAN	+	+	+	+	+	+	+	+
METROSIM	+	+	+	+	+	+		+
MUSSA	+	+			+	+		+
POLIS		+			+	+		+
RURBAN		+			+	+		+
STASA	+	+	+	+	+	+	+	+
TRANUS	+	+	+	+	+	+	+	+
URBANSIM		+	+	+	+	+		+

Source: [Wegener and Fürst, 2004]

Urban transportation models

MATSim model: heterogenous data and integration of many sub-models



Source: [Balmer et al., 2009]

MATSim model integration

Modular four-step multimodal transportation model using open source projects and data

Integrated models:

- MATSim model (MATSim Community) for the transportation system `https://www.matsim.org/` [W Axhausen et al., 2016]
- SPENSER model (University of Leeds) for the synthetic population `https://github.com/nismod/microsimulation` [Spooner et al., 2021]
- QUANT model (CASA, University College London) for spatial interactions to generate home-work plans `http://quant.casa.ucl.ac.uk/` [Batty and Milton, 2021]
- spatialdata library (OpenMOLE community) for data processing `https://github.com/openmole/spatialdata` [Raimbault et al., 2020]

Data and implementation

Data: Generic for any Functional Urban Area (GHSL [Florczyk et al., 2019]) or any arbitrary area in the UK: NOMIS census, OrdnanceSurvey roads, Traveline National Dataset for public transport

Workflow system for the integration of submodels and validation experiments: OpenMOLE software `https://openmole.org/` [Reuillon et al., 2013]

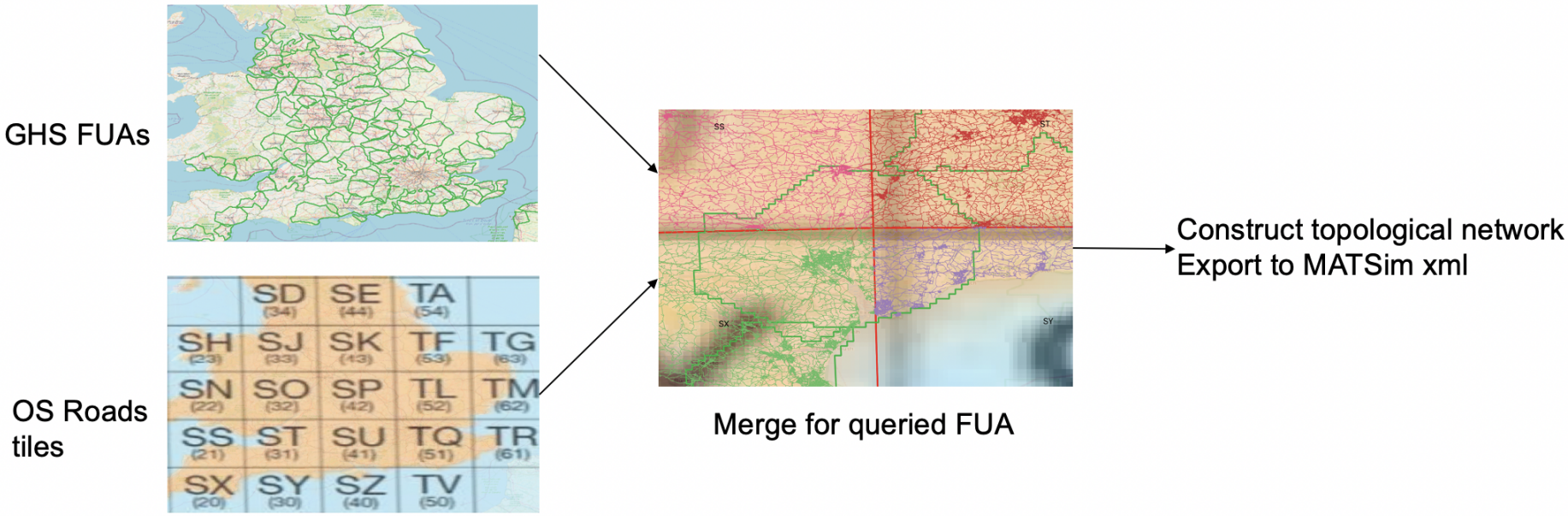
Implementation

- Synthetic SPENSER population distributed at the micro level using OSM buildings
- QUANT model to generate home-work commuting flows, job locations determined by sampling flows
- Network and plans (simple uniform commuting plans) prepared into MATSim xml files and fed into a multimodal MATSim model
- Models integrated as Docker containers

Open source: docker files to build model containers and integration script available at `https://github.com/JusteRaimbault/UrbanDynamics/tree/master/Models/Matsim`

Data preparation

→ Road network preprocessing: implemented into the `spatialdata` scala library [Raimbault et al., 2020]



→ Public transport data: from TransXchange (TNDS) to GTFS using UK2GTFS R package [Morgan, 2021]; GTFS to MATSim xml schedule using `pt2matsim` library

Explored parameters

Parameter sampled for the sensitivity analysis:

- Functional Urban Area (spatial context [Raimbault et al., 2019])
- Random seed (influence of stochasticity [Bienzeisler et al., 2021])
- Synthetic population sampling
- Modal choice parameters [Hörl, 2021]: mode constants in scoring function (car, public transport, walking)

OpenMOLE script

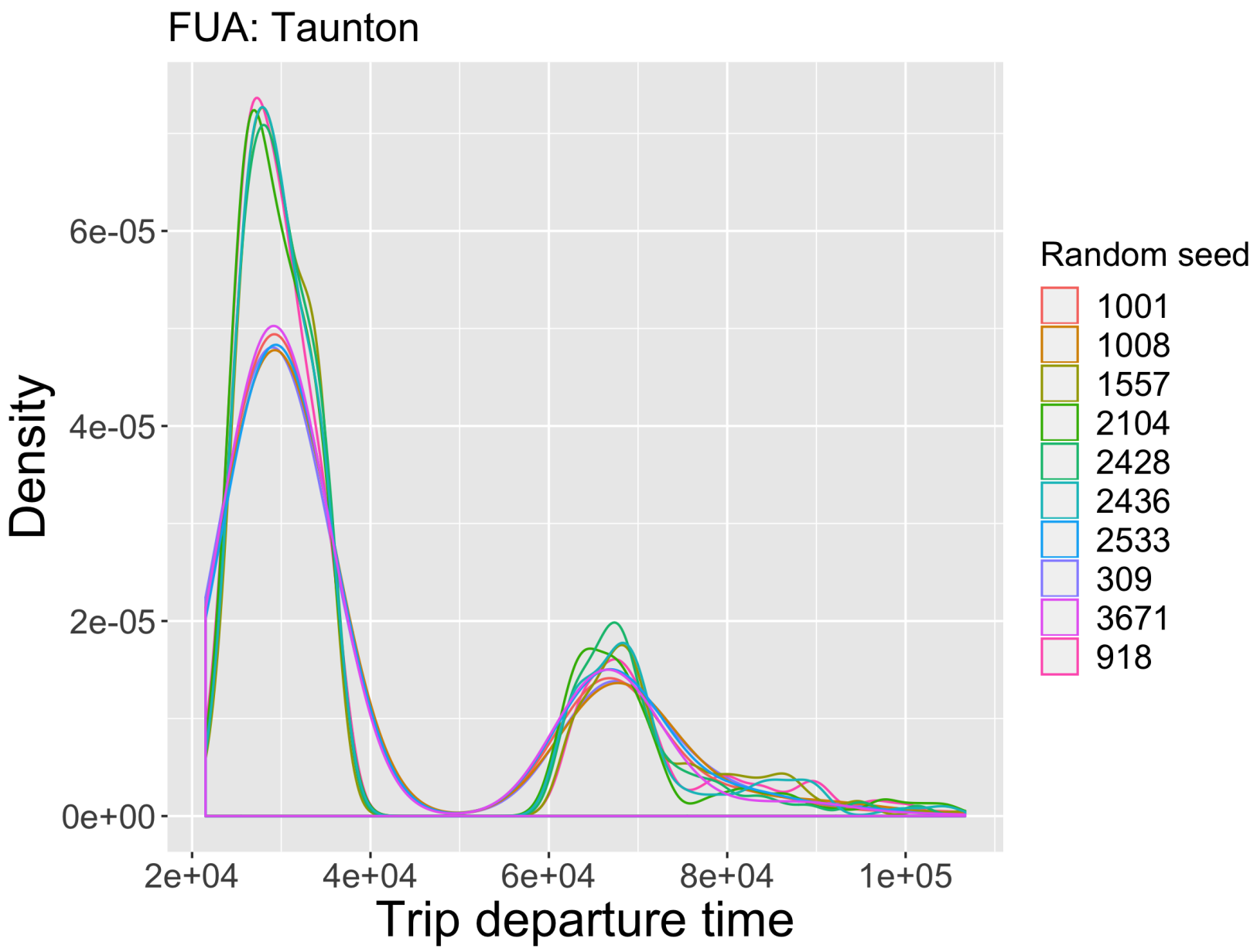
Saltelli.oms

```
val popSample = Val[Double]; val seed = Val[Double]; val fuaInd = Val[Double]; val utilitywalk = Val[Double]; val utilitycar = Val[Double]; val utilitypt = Val[Double]
val fuaName = Val[String];val fuaFile = Val[File]
val iterations = Val[Int]; val threads = Val[Int]
val dataDir = Val[File]; val nwDir = Val[File]
val planFile = Val[File]; val networkFile = Val[File]; val transitScheduleFile = Val[File]; val transitVehiclesFile = Val[File]
val outputDir = Val[File]
val carShare = Val[Double]; val ptShare = Val[Double]; val walkShare = Val[Double]; val avgTripDistance = Val[Double]; val avgScore = Val[Double]

val setFUA = ScalaTask(
  """
  val potFUAs = fuaFile.content.split("\n").toSeq
  val fuaName = potFUAs((fuaInd*potFUAs.size.toDouble).toInt)
  """
) set (
   (inputs,outputs) += (popSample,seed,utilitywalk,utilitycar,utilitypt),
   inputs += (fuaInd,fuaFile),
   outputs += (fuaName),
   fuaFile := workDirectory / "areas.txt"
)

val popcommand = "ls /data/LAD;export SBT_OPTS=\"-Xmx16G\";sbt \"project spatialdata; runMain org.openmole.spatialdata.application.matsim.RunMatsim --synthpop --popMode=detailed --jobMode=sample --planMode=default --sample=${popSample} --FUAName=${fuaName}"+
"--FUAFile=/data/GHS_FUA_UCDB2015_GLOBE_R2019A_54009_1K_V1_0_WGS84.gpkg --LAFile=/data/LAD/LAD_WGS84.shp --OAFile=/data/OA/OA2011_WGS84.shp --MSOAFile=/data/MSOA/EnglandWalesScotland_MSOAWGS84.shp --SPENSERDirs=/data/SPENSER/England,/data/SPENSER/Scotland,/data/SPENSER/
" --OSMBuildingsDirs=/data/OSM/England,/data/OSM/Scotland,/data/OSM/Wales --QUANTDataDir=/data/QUANT/ --output=Plans\";ls /data;[ ! -f Plans_${fuaName}.xml ] && touch Plans_${fuaName}.xml"

val population = ContainerTask(
    image = workDirectory / "images" / "matsim-population-1.0-0a52e7b80f.tar.gz",
    command = popcommand,
    workDirectory = "/root/spatialdata/library",
    containerSystem = Singularity(),
    errorOnReturnValue = false
) set (
   (inputs,outputs) += (fuaName,popSample,seed,utilitywalk,utilitycar,utilitypt),
   inputFiles += (dataDir,"/data/"),
   outputFiles += ("Plans_${fuaName}.xml",planFile),
   dataDir := workDirectory / "rawdata"
)

val readfiles = ScalaTask("""
   val networkFile = nwDir / (fuaName+".xml.gz")
   val transitScheduleFile = nwDir / (fuaName+"_transit_schedule.xml.gz")
   val transitVehiclesFile = nwDir / (fuaName+"_transit_vehicles.xml.gz")
"""
) set (
  (inputs,outputs) += (fuaName,popSample,seed,utilitywalk,utilitycar,utilitypt,planFile),
  inputs += (nwDir),
  outputs += (networkFile, transitScheduleFile, transitVehiclesFile),
  nwDir := workDirectory / "Network" / "runtime"
)
```

OpenMOLE script (continued)

```
51  val matsimcommand = """
52  export SEED=${seed.toInt};export ITERATIONS=${iterations};export THREADS=${threads}
53  export UTILITYWALK=${utilitywalk};export UTILITYPT=${utilitypt};export UTILITYCAR=${utilitycar}
54  ./config.sh
55  java -Xmx18G -cp matsim-12.0/matsim-12.0.jar org.matsim.run.Controler config_runtime.xml
56  mkdir output
57  """
58
59  val matsim = ContainerTask(
60    image = workDirectory / "images" / "matsim-1.0-86a2c28210.tar.gz",
61    command = matsimcommand,
62    workDirectory = "/root/",
63    containerSystem = Singularity(),
64    errorOnReturnValue = false
65  ) set (
66    (inputs,outputs) += (popSample,fuaName,seed,utilitywalk,utilitycar,utilitypt,iterations,threads),
67    inputFiles += (planFile,"Plans.xml"),
68    inputFiles += (networkFile,"Network.xml.gz"),
69    inputFiles += (transitScheduleFile,"transit_schedule.xml.gz"),
70    inputFiles += (transitVehiclesFile,"transit_vehicles.xml.gz"),
71    outputFiles += ("output",outputDir),
72    iterations := 1,
73    threads := 10,
74  )
75
76  val outputProcessing = ScalaTask("""
77    val modeShares: Array[Double] = if (outputDir.list.size>0) (outputDir / "modestats.txt").content.split("\n").last.split("\t").drop(1).map(_.toDouble) else Array(0.0,0.0,0.0)
78    val carShare = modeShares(0)
79    val ptShare = modeShares(1)
80    val walkShare = modeShares(2)
81    val avgTripDistance = if (outputDir.list.size>0) (outputDir / "traveldistancestats.txt").content.split("\n").last.split("\t").last.toDouble else 0.0
82    val avgScore = if (outputDir.list.size>0) (outputDir / "scorestats.txt").content.split("\n").last.split("\t").dropRight(1).last.toDouble else 0.0
83  """
84  ) set (
85    (inputs,outputs) += (popSample,fuaName,seed,utilitywalk,utilitycar,utilitypt,iterations,threads),
86    inputs += (outputDir),
87    outputs += (carShare, ptShare, walkShare, avgTripDistance, avgScore),
88    outputDir := workDirectory / "Matsim" / "output"
89  )
90
91
92  val purpose = "SALTELLI"
93  val ts = ((new java.text.SimpleDateFormat("yyyyMMdd_HHmmss")).format(new java.util.Date()))
94  val indics_hook = AppendToCSVFileHook(workDirectory / "saltelli" / (ts+"_"+purpose+".csv"),arrayOnRow = true)
95
96  val model = setFUA -- population -- readfiles -- matsim -- (outputProcessing hook indics_hook)
97
98  SensitivitySaltelli(
99    evaluation = model,
100   inputs = Seq(
101       utilitywalk in (0.0,100.0),
102       utilitypt in (0.0,100.0),
103       utilitycar in (0.0,100.0),
104       seed in (0.0,10000.0),
105       fuaInd in (0.0,0.9999999999),
106       popSample in (0.05,0.1)
107       ),
108   outputs = Seq(carShare, ptShare, walkShare, avgTripDistance, avgScore),
109   sample = 6
110 ) hook (workDirectory / "saltelli" / (ts+" "+purpose))
```

Role of stochasticity



Global Sensitivity Analysis

Method based on the estimation of conditional relative variances [Saltelli et al., 2010]

First order index

$$S_i = \frac{Var\left[E_{\mathbf{X}_{\sim i}}(Y|X_i)\right]}{Var\left[Y\right]}$$

is the expected relative variance reduction if X_i would be fixed

Total effect index

$$ST_i = \frac{E_{\mathbf{X}_{\sim i}}\left[Var\left(Y|\mathbf{X}_{\sim i}\right)\right]}{Var\left[Y\right]}$$

is the expected relative variance if all factors but X_i are fixed (includes interaction effects)

GSA results

output	β_W	β_{PT}	β_C	S	FUA	p
carShare	0.023	0.0058	0.0079	3.94	0.165	0.379
ptShare	0.0081	0.0074	0.0030	2.164	0.04	0.0169
walkShare	0.0059	0.0017	0.0074	0.834	0.16	0.082
avgTripDistance	0.11	0.19	0.087	0.04	1.51	0.049
avgScore	0.43	0.0003	0.0039	0.057	0.0085	0.0073

Total order Saltelli indices obtained with $\simeq 50$ *model runs*

Work in progress: Use OpenMOLE implementation of a genetic algorithm for multi-objective optimisation (NSGA2) to calibrate mode parameters on share observed for census commuting flows.

1. Integration of urban models
2. Modular open transport models
3. Anticipating the impact of future transportation infrastructure
4. Discussion

Coupling QUANT and SPENSER models

	QUANT	SPENSER
Time scale	10 years	40 years
Spatial scale	UK	UK
Spatial resolution	MSOA	MSOA
Agent granularity	Aggregated counts	individual level
Static/Dynamic	Equilibrium (static)	Dynamic
Randomness	Deterministic	Monte-Carlo
Transportation	3 modes	NA
Economics	Accessibility-based relocations	NA
Demographics	NA	Data-driven
Migration flows	Accessibility-based relocations	Data-driven

Weak coupling SPENSER (microsimulation) $\rightarrow$ QUANT (LUTI): $\simeq$ applying QUANT with future population $\simeq$ future impacts of a transportation infrastructure.

Current work and prospects

Current coupling: OpenMOLE script with containers for SPENSER (python/cpp) and the original QUANT backend (C#)

`https://github.com/JusteRaimbault/UrbanDynamics/blob/master/Models/SpenserQuantCoupling`

Work in progress: extension of QUANT implementation in spatialdata for more flexibility on scenario exploration, application to the CAMCOX corridor.

Open questions:

- Meaning of the weak coupling QUANT $\rightarrow$ SPENSER
- Stronger coupling: iteration of each model at each time step (progressive infrastructure construction, evolution of employments
- Application: genetic algorithm exploration of potential infrastructure, trade-offs accessibility/emissions

1. Integration of urban models
2. Modular open transport models
3. Anticipating the impact of future transportation infrastructure
4. Discussion

Discussion

$\rightarrow$ a theoretical and methodological framework to integrate model and associated tools - still does not solves all technical issues, model tuning, adaptation and optimisation may be necessary (e.g. MATSim: memory; QUANT-SPENSER: implementation). *Possibility of totally seamless model manipulation and coupling?*

$\rightarrow$ Clear definition of model coupling and its measure still needed: between scales, ontologies, (in)compatibilities.

$\rightarrow$ Validation of coupled models ($\neq$ coupling validated models).

$\rightarrow$ Role of model functions, which may change at integration [Varenne, 2018].

Conclusion

→ **Integrated models** capturing complexity of **urban sustainability** towards **decision-making**.

→ **Robust** knowledge from models obtained with the development of **validation and exploration methods**.

→ **Complementarity** of perspectives, models and theories in an **interdisciplinary** context.

To use OpenMOLE (free and open software) and contribute:
`https://openmole.org`

Presented models are open source at
`https://github.com/JusteRaimbault/UrbanDynamics`

References I

Balmer, M., Rieser, M., Meister, K., Charypar, D., Lefebvre, N., and Nagel, K. (2009).
Matsim-t: Architecture and simulation times.
In *Multi-agent systems for traffic and transportation engineering*, pages 57–78. IGI Global.

Banos, A. (2013).
Pour des pratiques de modélisation et de simulation libérées en géographie et SHS.
PhD thesis, Université Paris 1 Panthéon Sorbonne.

Batty, M. (2014).
Can it happen again? planning support, lee's requiem and the rise of the smart cities movement.
Environment and Planning B: Planning and Design, 41(3):388–391.

References II

Batty, M. and Milton, R. (2021).
A new framework for very large-scale urban modelling.
Urban Studies, 58(15):3071–3094.

Bienzeisler, L., T., L., Wage, O., Huck, L., and Friedrich, B. (2021).
Uncertainty and variability analysis of agent-based transport models.
EWGT2021 Proceedings.

Cottineau, C., Reuillon, R., Chapron, P., Rey-Coyrehourcq, S., and Pumain, D. (2015).
A modular modelling framework for hypotheses testing in the simulation of urbanisation.
Systems, 3(4):348–377.

References III

Florczyk, A. J., Corbane, C., Ehrlich, D., Freire, S., Kemper, T., Maffenini, L., Melchiorri, M., Pesaresi, M., Politis, P., Schiavina, M., et al. (2019).
Ghsl data package 2019.
Luxembourg, EUR, 29788(10.2760):290498.

Hörl, S. (2021).
Integrating discrete choice models with matsim scoring.
Procedia Computer Science, 184:704–711.

Lee Jr, D. B. (1973).
Requiem for large-scale models.
Journal of the American Institute of planners, 39(3):163–178.

Morgan, M. (2021).
UK2GTFS: Converts UK transport timetable datasets to GTFS format.
https://github.com/itsleeds/uk2gtfs, https://itsleeds.git.io/uk2gtfs/.

References IV

Raimbault, J. (2018).
Multi-modeling the morphogenesis of transportation networks.
In *European Conference on Artificial Life (ALIFE 2018)*, volume 2018, pages 382–383. MIT Press.

Raimbault, J. (2019a).
Exploration of an interdisciplinary scientific landscape.
Scientometrics, 119(2):617–641.

Raimbault, J. (2019b).
Second-order control of complex systems with correlated synthetic data.
Complex Adaptive Systems Modeling, 7(1):1–19.

Raimbault, J. (2020).
A comparison of simple models for urban morphogenesis.
arXiv preprint arXiv:2008.13277.

References V

Raimbault, J. (2021a).
Complementarity of generative models for road networks.
arXiv preprint arXiv:2109.15206.

Raimbault, J. (2021b).
Modeling the co-evolution of cities and networks.
In *Handbook of Cities and Networks*, pages 166–193. Edward Elgar Publishing.

Raimbault, J. (2021c).
A multiscale model of urban morphogenesis.
arXiv preprint arXiv:2103.17241.

Raimbault, J. (2021d).
Strong coupling between scales in a multi-scalar model of urban dynamics.
arXiv preprint arXiv:2101.12725.

References VI

Raimbault, J. and Batty, M. (2021).
Estimating public transport congestion in uk urban areas with open transport models.
arXiv preprint arXiv:2104.14359.

Raimbault, J., Cottineau, C., Le Texier, M., Le Néchet, F., and Reuillon, R. (2019).
Space matters: Extending sensitivity analysis to initial spatial conditions in geosimulation models.
Journal of Artificial Societies and Social Simulation, 22(4).

Raimbault, J. and Perret, J. (2019).
Generating urban morphologies at large scales.
In *The 2019 Conference on Artificial Life*, pages 179–186. MIT Press.

Raimbault, J., Perret, J., and Reuillon, R. (2020).
A scala library for spatial sensitivity analysis.
arXiv preprint arXiv:2007.10667.

References VII

Raimbault, J. and Pumain, D. (2022).
Trade-offs between sustainable development goals in systems of cities.
Journal of Urban Management.

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.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010).
Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index.
Computer physics communications, 181(2):259–270.

References VIII

Spooner, F., Abrams, J. F., Morrissey, K., Shaddick, G., Batty, M., Milton, R., Dennett, A., Lomax, N., Malleson, N., Nelissen, N., et al. (2021).
A dynamic microsimulation model for epidemics.
Social Science & Medicine, 291:114461.

Varenne, F. (2018).
Théories et modèles en sciences humaines: le cas de la géographie.
Éditions Matériologiques.

W Axhausen, K., Horni, A., and Nagel, K. (2016).
The multi-agent transport simulation MATSim.
Ubiquity Press.

Wegener, M. and Fürst, F. (2004).
Land-use transport interaction: State of the art.
Available at SSRN 1434678.

References IX

Zhuge, C., Shao, C., Wang, S., and Hu, Y. (2019).
Sensitivity analysis of integrated activity-based model: Using matsim as an example.
Transportation Letters, 11(2):93–103.