Models Coupling Urban Growth and Transportation Network Growth : An Algorithmic Systematic Review Approach

J. Raimbault[1,2]

juste.raimbault@polytechnique.edu

[1]UMR CNRS 8504 Géographie-cités
[2]UMR-T IFSTTAR 9403 LVMT

ECTQG 2015 Bari
Parallel Session : new data, methods
September 7, 2015

Lost in Translation

Science is all about *perspective* [Giere, 2010]...

Antagonist perspectives ?

G. Caruso saturday : Spatial ABM and Urban economy should (have to) reconciliate !
but even economist disagree [Marchionni, 2004] : geo. eco. *vs* eco. geo.

Listen to your teachers !
F. Leurent : your traffic model is wrong because out-of-eq.
M. Barthelemy : now we build a rigorous "urban science", all existing modeling work is garbage.

That escalated quickly... scientific aggressivity ! [Dupuy and Benguigui, 2015]
→ Do we "really stand together" ? [Pavlidis et al., 2014]

Research context

Thematic framework of the study of *transportation-cities (land-use) interactions*, on which misleading conceptions dominate [Offner, 1993].

[Bretagnolle, 2009] draws powerful empirical conclusions and claims for a need of **modeling the co-evolution process**.

$\rightarrow$ *State-of-the-art of such an hybrid modeling ?*

Disclaimer : NOT a thematic presentation on network/urban coevolution but more an applied epistemological study.

Brief Litterature Review

- LUTI approaches : transport integrated but static network (reviews : [Chang, 2006, Iacono et al., 2008] [Wegener and Fürst, 2004]). Various refinements (e.g. [Delons et al., 2008] housing market), many aspects can be taken into account [Wegener et al., 1991]
- Network Growth : Economic network growth [Zhang and Levinson, 2007, Xie and Levinson, 2009] ; Physical approaches [Barthélemy and Flammini, 2008] ; Biological self-organizing network [Tero et al., 2010].
- Hybrid dynamic approaches : few works propose dynamic coupled models, always toy-models. [Caruso et al., 2011, Moreno et al., 2009][Schmitt, 2014][Raimbault et al., 2014] [Le Néchet and Raimbault, 2015]

Research questions

Primal research question :

Why so few contributions on modeling this crucial problem ? What could be the role of the well-known disciplinar compartmentalization on that issue ?

Dual research question :

How does lack of pluridisciplinarity influences scientific outcomes ? Can this be understood through quantitative epistemology analyses ?

Work hypothesis

Possible Explanations :

- Too difficult problem ? (Fear of negative results [Ioannidis, 2005])
  "*A tale of complex coupling of spatio-temporal socio-technical processes at the good scales*" : Good luck !
- Tried before (a long time ago) and forget ?
  [Rey-Coyrehourcq, 2015]
- Compartmentalization of disciplines having different ontological frameworks, research objectives, etc.
  [Commenges, 2013]

→ Work on testing quantitatively the last hypothesis.

Proposed approach

Bibliometric quantitative analysis

- Bibliometry and Systematic Reviews (SR) play nowadays a significant role in many disciplines (see SR/Meta-Analyses in therapeutic evaluation [Rucker, 2012]) ; 11 SR.day^{-1} [Bastian et al., 2010] !
- SR may even be considered as mandatory as missing a reference is more a mistake given new ressources [Lissack, 2013].
- Interesting results such as models predicting success of a paper and scientific success [Wang et al., 2013]
- Current bibliometric network study do rely on citation networks [Newman, 2013] or co-autorship networks [Sarigöl et al., 20

Proposed approach

$\rightarrow$ *Iterative construction of semantic neighborhoods in scientific reference network.*

- Extension of the method proposed in [Chavalarias and Cointet, 20 where a dynamic cartography of scientific disciplines is built through text-mining (*phylomemy*).
- Based on generalized keywords extraction from titles and abstracts of references, we propose to iteratively explore the keyword co-occurence network.
- If disciplines are indeed "closed", we should converge towards communities.

Algorithm formalization

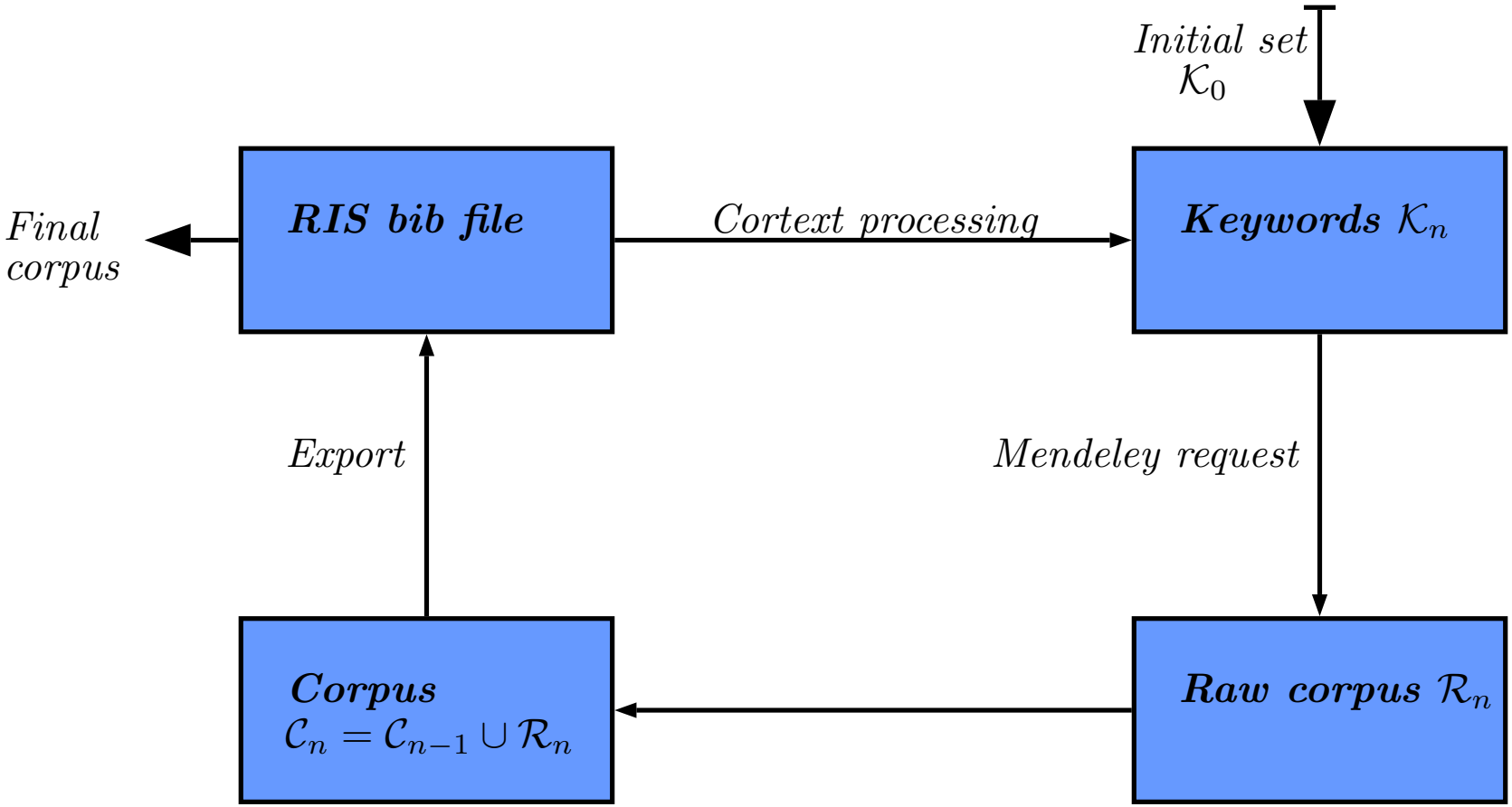
Let $\mathscr{A}$ an alphabet, $\mathscr{A}^*$ words and $\mathscr{T} = \cup_{k \in \mathbb{N}} \mathscr{A}^{*k}$ finite texts on it ; set of references at iteration n denoted $\mathscr{C}_n \in \mathscr{T}^3$. Set of keywords $\mathscr{K}_n$, initial keywords $\mathscr{K}_0$

Iteration of the algorithm

- Intermediate corpus $\mathscr{R}_n$ is obtained through a catalog request from previous keywords $\mathscr{K}_{n-1}$.
- New corpus : $\mathscr{C}_n = \mathscr{C}_{n-1} \cup \mathscr{R}_n$.
- New keywords $\mathscr{K}_n$ extracted from corpus through Natural Language Processing, with N_k fixed number of keywords.

Stopping criteria : converged corpus size or fixed max iteration

Algorithm architecture



Convergence : sketch of proof

1st attempt : we have obviously, with K, C keyword extraction and catalog request functions, $|\mathscr{C}_n| \leq |\mathscr{C}_{n+1}| \leq |\mathscr{C}_n| + |C[K(\mathscr{C}_n)]|$, what gives $|C[K(\mathscr{C}_n)]| \to 0 \iff |\mathscr{C}_n|$ bounded $\iff \mathscr{C}_n$ converges. *But assumptions difficultly interpretable*

2nd attempt :

- Assumption 1 : existence of *communities* $\mathscr{S}$ in network, on which K is stable.
- Assumption 2 : *random* steering behavior of the corpus (if not, converges as it growths)
- Then $T = \inf_n \mathscr{C}_n \in \mathscr{S}$ is a Stopping Time and thus convergence by stability of any $S \in \mathscr{S}$.

Implementation

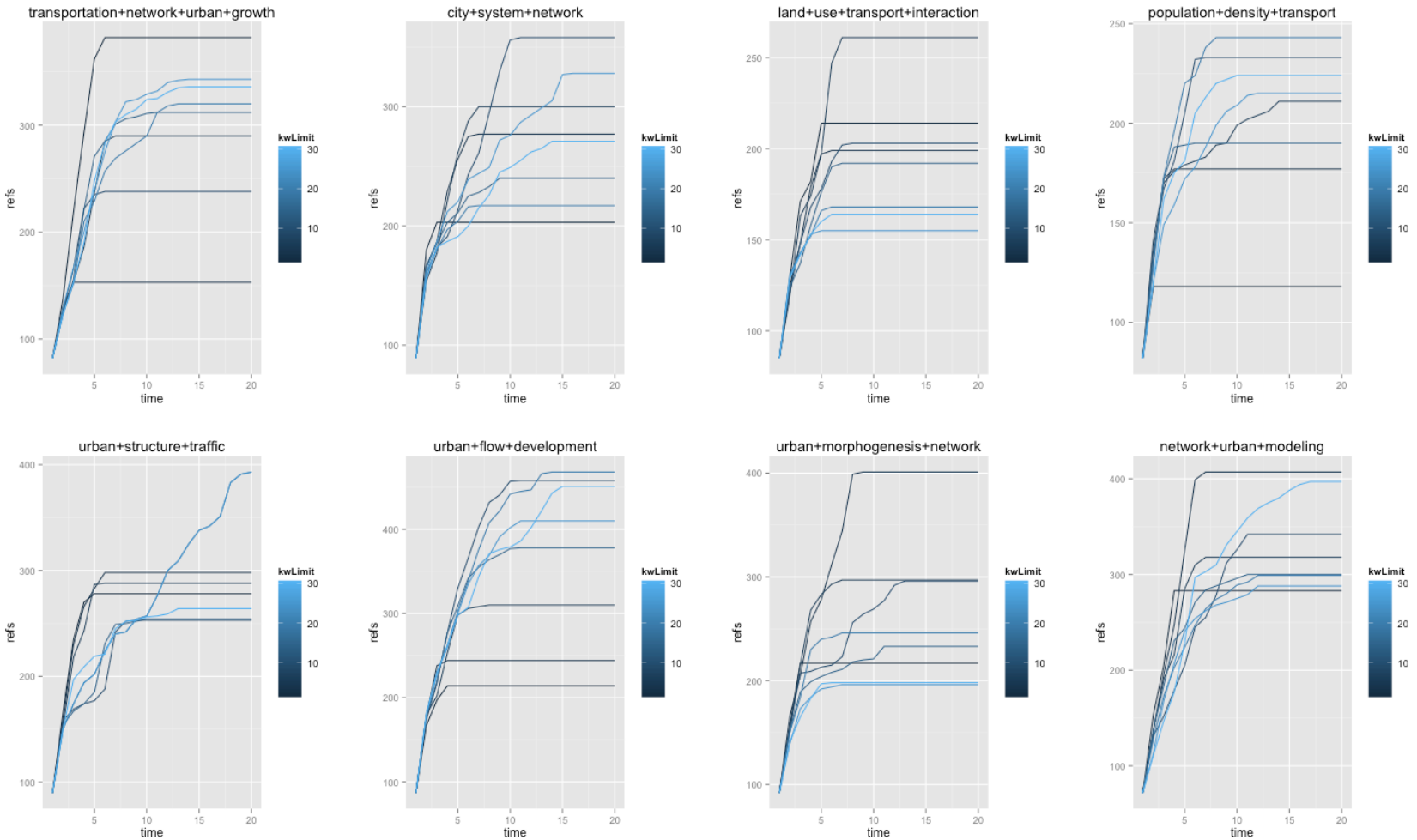
$\rightarrow$*Heterogeneity of functional specifications suggest an hybrid architecture (applied language agnosticity !) :*

- Core implemented in java ; experiment module in R
- Catalog request done to Mendeley open catalog [Mendeley, 2015] through REST API.
- Natural Language Processing through an hacked API of the web platform provided openly by [Chavalarias and Cointet, 2013] in the frame of the *Cortext* project.

Openly available at
https://github.com/JusteRaimbault/CityNetwork/tree/master/Models/Biblio/AlgoSR/AlgoSRJavaApp

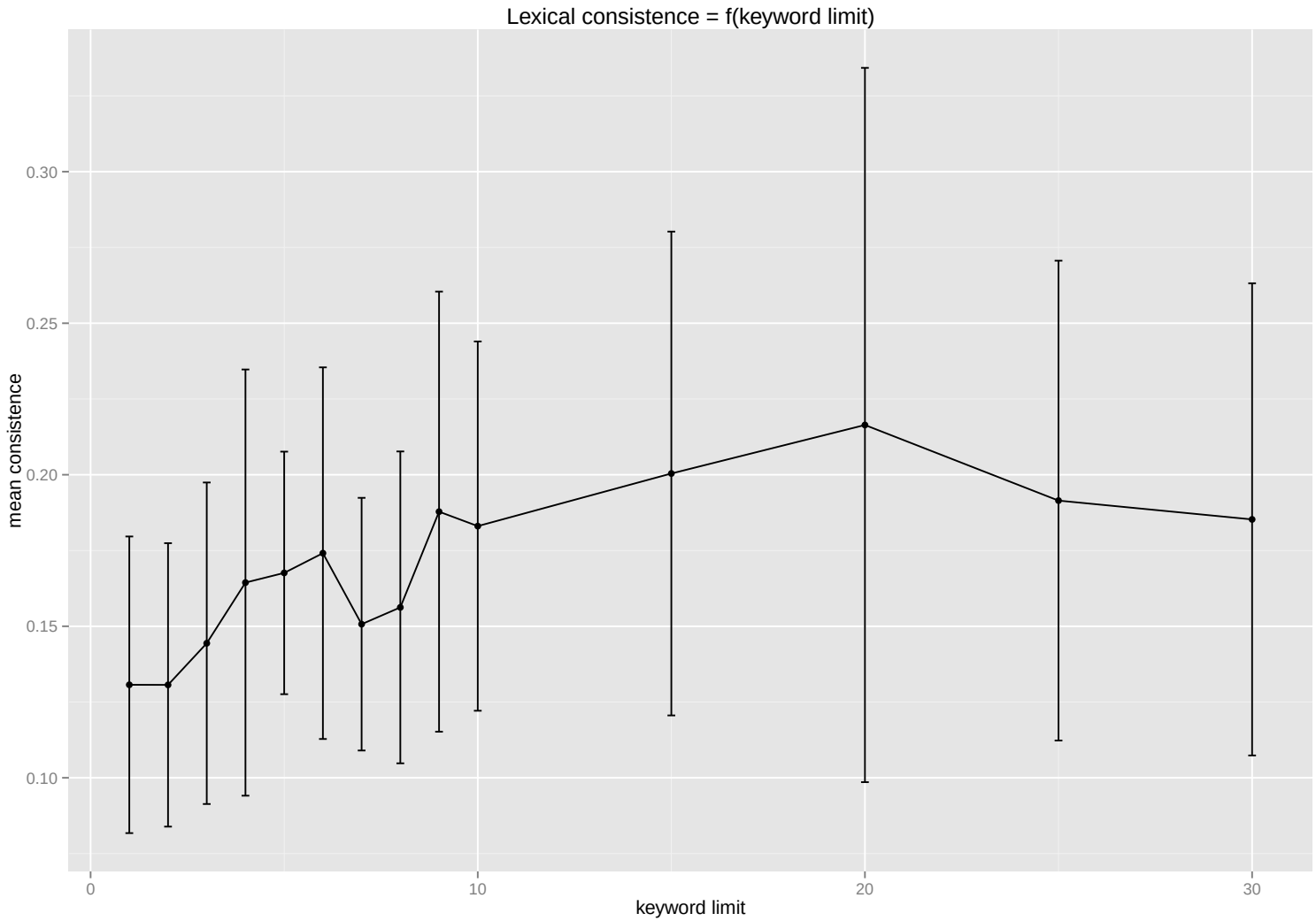
Algorithm convergence

Influence of parameters on convergence behavior



Algorithm consistence

Corpus lexical consistence



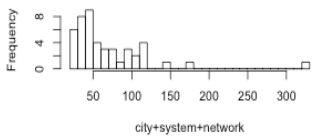
Application : experimental design

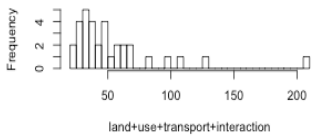
→ **Initial entries :** `city+system+network`, `land+use+transport+interaction`, `network+urban+modeling`, `population+density+transport`, `transportation+network+urban+growth`

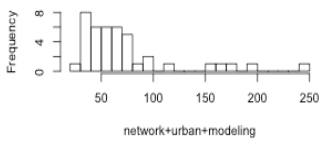
→ **Not restraining propagation parameter :** $N_k = 30$

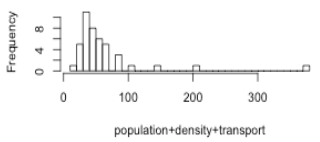
→ **Large convergence time** (obtained convergence steps : 16, 8, 17, 14,14)

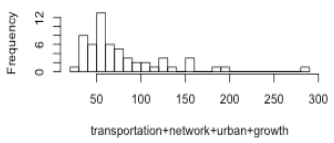
Application : corpuses distances











Corpuses	1	2	3	4	5
1 (W=3789)	1	0	0.0719	0.0078	0.0724
2 (W=5180)	0	1	0.0338	0	0.0125
3 (W=3757)	0.0719	0.0338	1	0.0100	0.1729
4 (W=3551)	0.0078	0	0.0100	1	0.0333
5 (W=8338)	0.0724	0.0125	0.1729	0.0333	1

Application : ongoing extensions

1. Method of construction of controlled random synthetic corpuses, to construct a kind of *null model* to have objective benchmark on intercluster rates.

2. Comparison of results obtained by semantic analysis with results from citation network or co-authorship network ; should find similar patterns.
   Problem : Construction of citation network from **Open Data** is tricky ; elaboration of a scholar API avoiding google blocking using a TOR threads pool.

Discussion

- Quality of / robustness to catalog [Bohannon, 2014] has to be more investigated.
- "Subjective" initial review and corpuses choices should be more systematically based on expert knowledge.
- More systematic analysis and exploration of behavior, as for a computationnal model (concept of proof in Social Sciences ? [Banos, 2013][Laughlin, 2006]
- Need to use/implement analysis and visualization tools such as in [Cuyala and Commenges, 2015]

Perspectives

- Project of a qualitative work (interviews of "old" experts, archives, grey litterature investigation) following the methodology of [Commenges, 2013].
- Algorithm as an Open source software available, easy to use by anybody $\Longrightarrow$ perspective of data collection on algorithm behavior ; collaborative building of local science maps from expert knowledge.
- Transposition of the methodology in other bibliometrical projects.

Conclusion : Proposition and implementation of an iterative semantic discipline mapping algorithm ; first results suggest a positive answer to the compartmentalization hypothesis.

Reserve slides

Reserve Slides

Lexical consistence

Lexical consistence is defined though co-occurrences of keywords by, with N final number of keywords, f final step, and $c(i)$ co-occurrences in references,

$$\kappa = \frac{2}{N(N-1)} \sum_{i,j \in \mathscr{K}_f} |c(i) - c(j)|$$

Corpuses Distances

Corpus distances defined as

$$d(i,j) = \frac{\sum_{k,l} \mathbb{1}_{w_{ki}=w_{lj}} \cdot (c_{ki} + c_{lj})}{\sum_k (c_{ki} + c_{kj})}$$

References I


Banos, A. (2013).
Pour des pratiques de modélisation et de simulation libérées en géographies et shs.
HDR. Université Paris, 1.


Barthélemy, M. and Flammini, A. (2008).
Modeling urban street patterns.
Physical review letters, 100(13):138702.


Bastian, H., Glasziou, P., and Chalmers, I. (2010).
Seventy-five trials and eleven systematic reviews a day: How will we ever keep up?
PLoS Med, 7(9):e1000326.

References II


Bohannon, J. (2014).
Scientific publishing. google scholar wins raves–but can it be trusted?
Science (New York, NY), 343(6166):14.


Bretagnolle, A. (2009).
Villes et réseaux de transport : des interactions dans la longue durée, France, Europe, États-Unis.
Hdr, Université Panthéon-Sorbonne - Paris I.

References III


Caruso, G., Vuidel, G., Cavailhes, J., Frankhauser, P., Peeters, D., and Thomas, I. (2011).
Morphological similarities between dbm and a microeconomic model of sprawl.
Journal of Geographical Systems, 13:31–48.


Chang, J. S. (2006).
Models of the relationship between transport and land-use: A review.
Transport Reviews, 26(3):325–350.

References IV


Chavalarias, D. and Cointet, J.-P. (2013).
Phylomemetic patterns in science evolution—the rise and fall of scientific fields.
Plos One, 8(2):e54847.


Commenges, H. (2013).
The invention of daily mobility. Performative aspects of the instruments of economics of transportation.
Theses, Université Paris-Diderot - Paris VII.

References V


Cuyala, S. and Commenges, H. (2015).
The geographer's words: an interactive platform to study the evolution of the discipline.
In *Plurimondi. An International Forum for Research and Debate on Human Settlements*, volume 7.


Delons, J., Coulombel, N., and Leurent, F. (2008).
PIRANDELLO an integrated transport and land-use model for the Paris area.


Dupuy, G. and Benguigui, L. G. (2015).
Sciences urbaines: interdisciplinarités passive, naïve, transitive, offensive.
Métropoles, (16).

References VI


Giere, R. N. (2010).
Scientific perspectivism.
University of Chicago Press.


Iacono, M., Levinson, D., and El-Geneidy, A. (2008).
Models of transportation and land use change: a guide to the territory.
Journal of Planning Literature, 22(4):323–340.


Ioannidis, J. P. (2005).
Why most published research findings are false.
Chance, 18(4):40–47.

References VII


Laughlin, R. B. (2006).
A different universe: Reinventing physics from the bottom down.
Basic Books.


Le Néchet, F. and Raimbault, J. (2015).
Modeling the emergence of metropolitan transport autorithy in a polycentric urban region.
Plurimondi. An International Forum for Research and Debate on Human Settlements, 7(15).

References VIII


Lissack, M. (2013).
Subliminal influence or plagiarism by negligence ? the slodderwetenschap of ignoring the internet.
Journal of Academic Ethics.


Marchionni, C. (2004).
Geographical economics versus economic geography: towards a clarification of the dispute.
Environment and Planning A, 36(10):1737–1753.


Mendeley (2015).
Mendeley reference manager.
http://www.mendeley.com/.

References IX


Moreno, D., Badariotti, D., and Banos, A. (2009).
Un automate cellulaire pour expérimenter les effets de la proximité dans le processus d'étalement urbain : le modèle raumulus.
Cybergeo : European Journal of Geography.


Newman, M. E. J. (2013).
Prediction of highly cited papers.
ArXiv e-prints.


Offner, J.-M. (1993).
Les "effets structurants" du transport: mythe politique, mystification scientifique.
Espace géographique, 22(3):233–242.

References X


Pavlidis, I., Petersen, A. M., and Semendeferi, I. (2014).
Together we stand.
Nature Physics, 10(10):700–702.


Raimbault, J., Banos, A., and Doursat, R. (2014).
A hybrid network/grid model of urban morphogenesis and optimization.
In *Proceedings of the 4th International Conference on Complex Systems and Applications (ICCSA 2014), June 23-26, 2014, Université de Normandie, Le Havre, France; M. A. Aziz-Alaoui, C. Bertelle, X. Z. Liu, D. Olivier, eds.: pp. 51-60.*

References XI


Rey-Coyrehourcq, S. (2015).
These de Doctorat.
PhD thesis, UMR CNRS 8504 Geographie-cites.


Rucker, G. (2012).
Network meta-analysis, electrical networks and graph theory.
Research Synthesis Methods, 3(4):312–324.


Sarigöl, E., Pfitzner, R., Scholtes, I., Garas, A., and Schweitzer, F. (2014).
Predicting Scientific Success Based on Coauthorship Networks.
ArXiv e-prints.

References XII


Schmitt, C. (2014).
Modélisation de la dynamique des systèmes de peuplement: de SimpopLocal à SimpopNet.
PhD thesis, Paris 1.


Tero, A., Takagi, S., Saigusa, T., Ito, K., Bebber, D. P., Fricker, M. D., Yumiki, K., Kobayashi, R., and Nakagaki, T. (2010).
Rules for biologically inspired adaptive network design.
Science, 327(5964):439–442.


Wang, D., Song, C., and Barabási, A.-L. (2013).
Quantifying long-term scientific impact.
Science, 342(6154):127–132.

References XIII


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


Wegener, M., Mackett, R. L., and Simmonds, D. C. (1991).
One city, three models: comparison of land-use/transport policy simulation models for dortmund.
Transport Reviews, 11(2):107–129.


Xie, F. and Levinson, D. (2009).
Modeling the growth of transportation networks: A comprehensive review.
Networks and Spatial Economics, 9(3):291–307.

References XIV


Zhang, L. and Levinson, D. (2007).
The economics of transportation network growth.
In *Essays on transport economics*, pages 317–339. Springer.