

For a Cautious Use of Big Data and Computation

Communication Proposal, RGS 2016 *Session : GeoComputation ; the next 20 years*

JUSTE RAIMBAULT^{1,2}

(1) UMR CNRS 8504 Géographie-cités and (2) UMR-T IFSTTAR 9403 LVMT

Keywords : *Data Deluge, Computational Science, City-transportation Interactions, Spatio-temporal Correlations*

The so-called *big data revolution* resides as much in the availability of large datasets of novel and various types as in the always increasing available computational power. Although the *computational shift* (Arthur (2015)) is central for a science aware of complexity and is undeniably the basis of future modeling practices in geography as Banos (2013) points out, we argue that both *data deluge* and *computational potentialities* are dangerous if not framed into a proper theoretical and formal framework. The first may bias research directions towards available datasets (as e.g. numerous twitter mobility studies) with the risk to disconnect from a theoretical background, whereas the second may overshadow preliminaries analytical resolutions essential for a consistent use of simulations. We illustrate this idea with an example on well-studied geographical objects that are interactions between networks and territories. We compute static correlations between indicators of urban form and indicators of road network topology, using open datasets of european population density (EUROSTAT (2014)) and OpenStreetMap. A mathematical derivation of the link between spatial covariance at fixed time and dynamical covariance for spatio-temporal stochastic processes, combined with a theory of city-transportation interactions within evolutive urban systems on long times (Bretagnolle (2009)), allows to infer knowledge on involved *geographical processes* from empirical static correlations. In particular we show the regional nature of dynamic interactions, confirming the non-ergodicity of urban systems (Pumain (2012)). We argue that the two conditions for this result are indeed the ones endangered by incautious big-data enthusiasm, concluding that a main challenge for future Geocomputation is a wise integration of novel practices within the existing body of knowledge.

References

- Arthur, W. B. (2015). Complexity and the shift in modern science. Conference on Complex Systems, Tempe, Arizona.
- Banos, A. (2013). Pour des pratiques de modélisation et de simulation libérées en géographies et shs. *HDR. Université Paris*, 1.
- Bretagnolle, A. (2009). *Villes et réseaux de transport : des interactions dans la longue durée, France, Europe, États-Unis*. Hdr, Université Panthéon-Sorbonne - Paris I.
- EUROSTAT (2014). Eurostat geographical data.
- Pumain, D. (2012). Urban systems dynamics, urban growth and scaling laws: The question of ergodicity. In *Complexity Theories of Cities Have Come of Age*, pages 91–103. Springer.