# Spatial Data Science Needs Explanations!
## (And Not Just Predictions)

**Mark Gahegan**
The University of Auckland, New Zealand (m.gahegan@auckland.ac.nz)

## Introduction

Spatial Data Science facilitates new approaches to longstanding problems in health, policy, ecology and economy. But to make the most effective use of it, we need analytical methods that are straightforward to apply, open to review and audit, and produce results that can be correctly interpreted by practicing researchers and policymakers.

Leveraging data science presents many challenges that manifest in unique ways in the geospatial realm, including: the nature of errors and bias in data, the opportunities afforded by new and dynamic data sources, the need for transparency and the ability to question assumptions where outcomes are used for public policy. The model-free nature of popular data science methods means they typically operate as black boxes. As such it is difficult to validate their outcomes in ways that stand up to policy debates and foster public trust.

The capacity of data science to help us *explain* and *understand* the underlying causes for the patterns we find in data has lagged behind the development of new analytical methods (Shmueli, 2010). But better explanation and transparency are necessary to empower end-users to interpret outputs in ways that conceptually align with pre-existing knowledge (Doshi-Velez, and Kim, 2017; Lipton, 2016).

From this perspective, there is currently an over-dependence on correlative models that neither represent process nor embed theory (O'Sullivan, 2018). The most capable predictive methods are typically opaque, with no explanation, and no error distributions that connect back to the underlying mechanisms that are driving outcomes. This limits the insights obtained, and our ability to understand the process and causal interactions in the underlying system. By contrast, the focus here is on *explanation*, via methods that strive towards *abductive* reasoning that can offer an explanation (hypothesis) for a set of observations (Gahegan, 2009).

Moving from prediction to explanation, specifically explanation based on feasible processes and their interactions, represents a significant and important step for the GIScience. Models will then become useful in terms of their predictive power, **and** their ability to explain likely processes and mechanisms that generate those predictions.

## Searching for Explanations

Inductive Process Modelling is an example of one kind of approach that can create explanatory models from complex data. It does this using a kind of theory and data *soup*: the soup contains a set of theory fragments, rules, constraints, and observational data drawn from the system under investigation. Using these building blocks, heuristic search and feedback are used to iteratively assemble pieces together into a chain that links observable variables with each other via causal mechanisms. Since the model produced contains a set of equations with known links to current theory, then it becomes possible to 'read' the model; it is essentially a hypothesised mathematical description of the process under investigation. See Asgharbeygi *et al.* (2006),

Sozou *et al.* (2017) and Gahegan (2019) for more details. When coupled with the descriptive richness of big data, this approach can produce complex and accurate models that can be readily understood, debated and validated.

The inductive process modelling community has had good success in constructing models from complex data that describe various spatial and temporal processes. As Langley showed nearly forty years ago, it is possible to 'discover' some well-known physical laws from data, such as acceleration due to gravity (Langley, 1981). In this case, the model is simple: just one equation with one rate term, and it can indeed be learned autonomously from experimental data. But more recent work includes discovering ecological models that can describe: complex predator-prey relationships (Bridewell et al, 2008), the occurrence of large algal blooms (Džeroski *et al.*, 2007) and complex reaction networks in systems biology (Džeroski & Todorovski, 2008). All of these examples not only demonstrate excellent results—they are every bit as good as those produced by expert scientists, and sometimes better—but they are also directly comparable with existing models and their correctness (or otherwise) can be debated rationally based on theory, not just on producing the best results following extensive training. Recent work by Langley and Arvay (2016) demonstrates how Inductive Process Modelling can scale to very complex ecological food webs with 15 or more different interacting species, all successfully learned from data describing the observed counts of individual organisms. That's a very complex model, with an intricate and lagged set of co-dependencies.

To give a specifically geographical example of model discovery, imagine that we wish to understand the mechanisms that drive overland water drainage and soil erosion. Some simple examples of equations to add to the process library might be: exponential decay, simple geometric calculations of upslope area, water accumulation, soil erosion coefficients and equations governing fluid flow and saturation. To guide the model discovery process we need a measure of success, which can be straightforwardly defined as a reduction of error in the model's ability to correctly predict the observed data (say the outflow of the basin or the amount of soil eroded over time). This provides the feedback needed to guide and refine the model's development. We will also need observation data that describes the variables that appear in the equations in the process library.

At the core of the computational model discovery movement is the idea that not only can data be turned into knowledge, but that this knowledge can in turn be expressed as theory (Horacek, 2017; Rule *et al.*, 2018). When the data mining or machine learning community refers to knowledge discovery, it typically means uncovering the patterns and categories that characterize trends or outliers in data; such 'knowledge' is usually expressed in the language of data (statistics) or machine learning method parameterisations and *not* the language of human explanation (domain theory). But in the case of computational model discovery, the discovered knowledge is a mathematical model (e.g. a set of Partial Differential Equations) that characterises some phenomena of interest and that can be translated (via these equations) into rules and relations between data—for example as processes and causal links. This is closer to theory, and can be interpreted from a theoretical perspective in many cases. Over time, as additional layers of the science process are adequately represented in computational systems, these discovery systems will become capable of producing rich and detailed descriptions of solutions and their explanations, couched in relevant domain theory and the process of research.

# References

Arvay, A. and Langley, P. (2016). Selective induction of rate-based process models. *Advances in Cognitive Systems*, 4, 1-15.

Asgharbeygi, N., Langley, P., Bay, S. and Arrigo, K. (2006). Inductive revision of quantitative process models, *Ecological Modelling* 194 (1-3), 70-79.

Bridewell, W., Langley, P., Todorovski, L., Džerosk, S. (2008) Inductive process modeling, *Machine learning* 71 (1), 1-32.

Doshi-Velez, F., and Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608.*

Džeroski, S., Langley, P., and Todorovski, L. (2007). Computational discovery of scientific knowledge. In S. Dzeroski and L. Todorovski (Eds.), *Computational discovery of communicable scientific knowledge.* Berlin: Springer.

Džeroski, S., & Todorovski, L. (2008). Equation discovery for systems biology: finding the structure and dynamics of biological networks from time course data. *Current Opinion in Biotechnology*, *19*(4), 360-368.

Gahegan, M. (2009). Visual exploration and explanation in geography: analysis with light. In: (Miller, H and Han, J., Eds) *Geographic data mining and knowledge discovery*, Ch 11, pp. 291-315.

Gahegan, M. (2019). Fourth Paradigm GIScience? Prospects for Automated Discovery and Explanation from Data. Forthcoming in International journal of Geographical Information Science.

Horacek, H. (2017). Requirements for Conceptual Representations of Explanations and How Reasoning Systems Can Serve Them. In *Proceedings of the 1st Workshop on Explainable Computational Intelligence (XCI 2017).*

Langley, P. (1981). Data-driven discovery of physical laws. *Cognitive Science*, 5, 31-54.

Lipton, Z. C. (2016) The mythos of model interpretability. *Queue* 16(3): 30:31--30:57.

O'Sullivan, D. (2018) Big data … why (oh why?) this computational social science? In: J. E. Thatcher, J. Eckert and A. Shears (eds), *Thinking Big Data in Geography: New Regimes, New Research*, (Lincoln: University of Nebraska Press), pages 21-38.

Rule, A., Tabard, A., and Hollan, J. D. (2018). Exploration and explanation in computational notebooks. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (p. 32). ACM.

Shmueli, G. (2010) To explain or to predict? *Statistical science* 25(3): 289-310.

Sozou P.D., Lane P.C., Addis M. and Gobet F. (2017). Computational Scientific Discovery. In: Magnani L., Bertolotti T. (eds.) *Springer Handbook of Model-Based Science.* Springer Handbooks. Springer, Cham.