What makes a system ‘complex’?

Identifying kinds of complex system and how to build them

Gary Polhill
The James Hutton Institute, Craigiebuckler, Aberdeen, UK
Email: gary.polhill@hutton.ac.uk

Introduction

‘Complex’ is a term used to describe various systems that are not amenable to study using traditional methods. Several observable features of systems make us suspect that they might be complex. Not least of these is leptokurtic (or ’fat-tailed’) distributions of sizes of event or other phenomena, meaning that parametric statistics (for which there is a significant body of mathematics) are unsuitable for analysing them. Heteroskedastic time series (the variance is not constant, and does not converge by increasing the length of the sample) are another phenomenon challenging for traditional mathematical analysis that may also indicate complexity. Similarly, spatial autocorrelation and statistical properties of networks may be indicative.

Equally, when stating that a system is ‘complex’, people may refer to ontological features of the system. Such features might be multiple heterogeneous interacting ‘parts’, nonlinear behaviour of those parts, partially interacting layers of structure in the system, absence of a global controller, or emergent system-level behaviour difficult to describe in terms of the behaviours of the parts.

Though some researchers aspire to a *general* theory of systems described as ‘complex’, this need not be the case. (See also Nick Gotts’s poster / presentation.) *Why shouldn’t there be numerous classes of complex system?* There are several classes of computational complexity. *There seems no logical reason to believe that all systems described as complex necessarily share common properties and are amenable to a single general theory*.

Acknowledgements
The work shown here was developed in collaboration with Andrew Jarvis (University of Lancaster), Nick Gotts (Independent Researcher), Volker Grimm (UFZ Leipzig), Laszlo Gulyas (AITIA), Andrew Adamatzy (University of the West of England) and Guillaume Deffaunt (Irstea).

Scottish Government
Riaghaltas na h-Alba
gov.scot

Method

If we suspect that there may be multiple classes of complex system then we need to find ways of discovering them. One approach is to look at some of the indicators of complexity that people use, such as time series, network and spatial statistics, and see whether combinations of these form distinct patterns. At the same time, we can examine the ontological features of the system that we think might cause complex dynamics, and see how these match to patterns in the spatial statistics we observe. Depending on whether we see a match, we can determine whether there might be a new ontological feature we need to examine, or whether we should be including different indicators in our analysis.

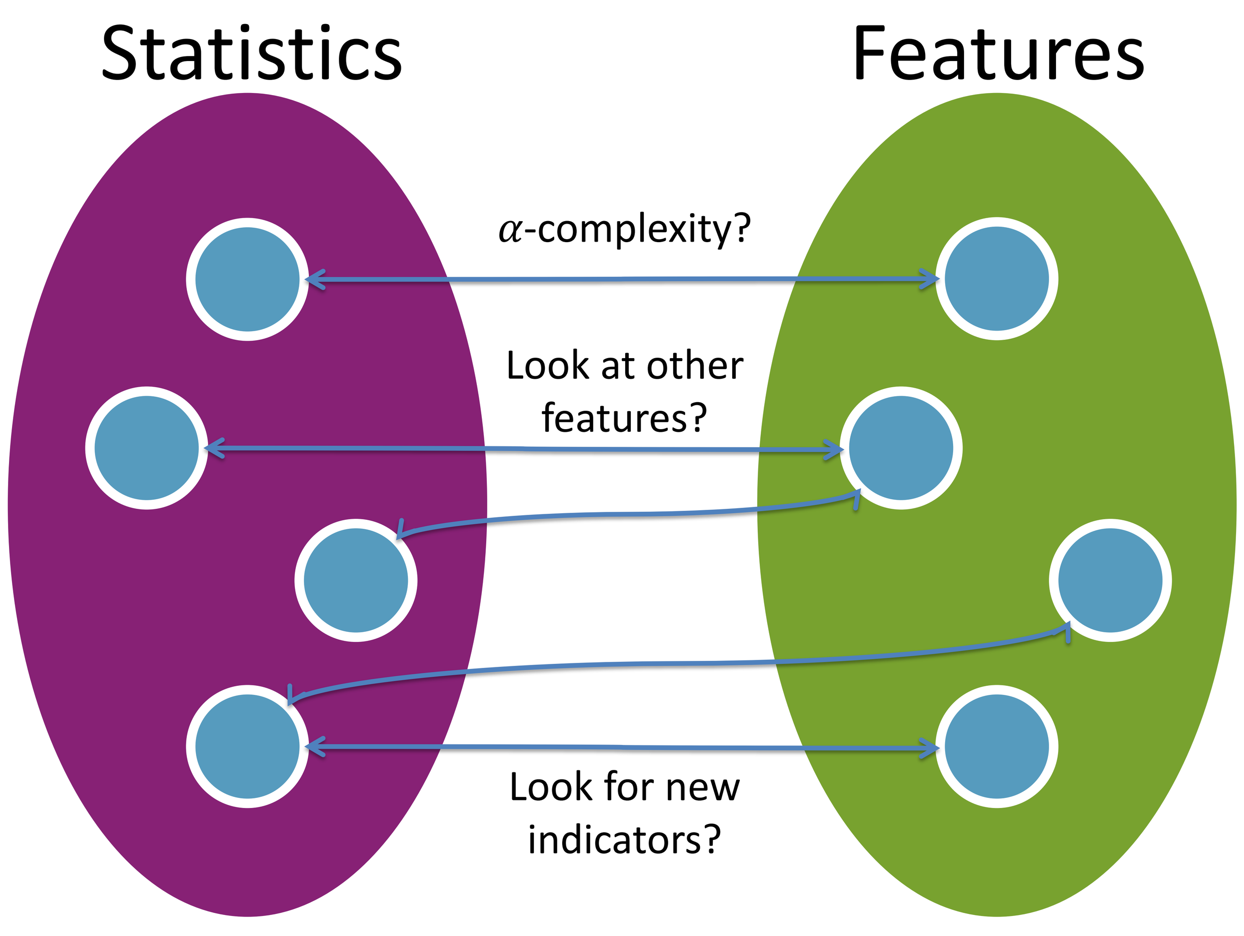

Figure 1 A method for finding classes of complex system by matching clusters of features and statistics

Explorations with FEARLUS-SPOMM

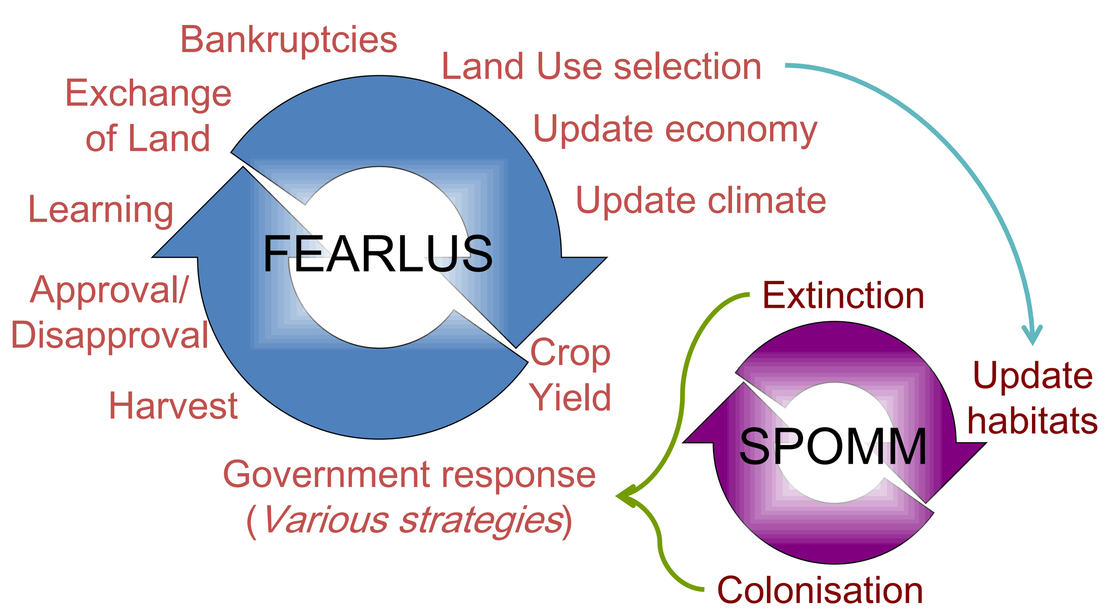

Figure 2 The FEARLUS-SPOMM model. Different feedback loops are created through the government’s response, which can be activity or outcome based, and include (or not) spatially clustered incentives.

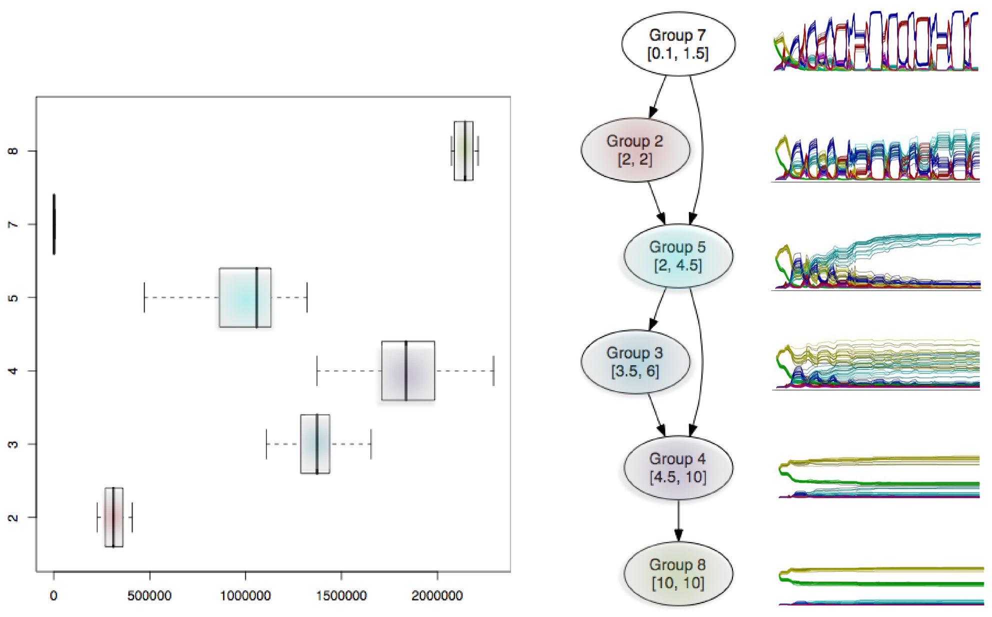

Figure 3 Clustering time series shows distinct patterns of behaviour with increasing incentive

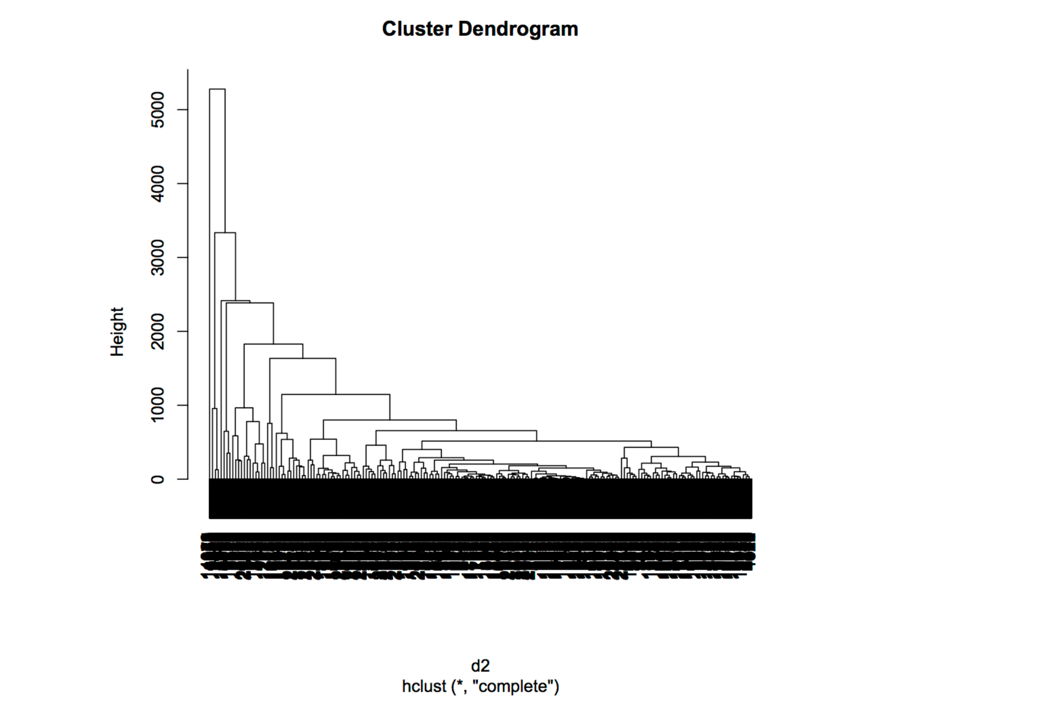

Figure 4 Clustering using various time series metrics (kurtosis, Hill tail exponent, ACF variance), over a number of variables and variations of the time series – 56 statistics in total for each of ~20,000 runs

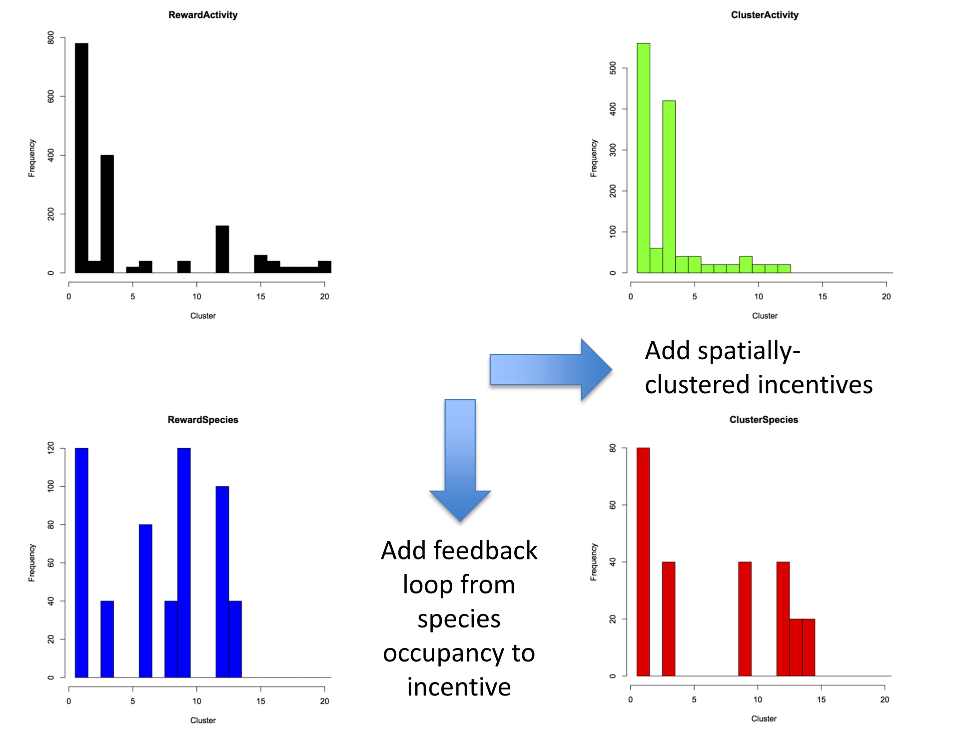

Figure 5 Mapping the clusters to ontological features

Observations

- These preliminary results do seem to suggest there may be distinct patterns of behaviour in the time series data of FEARLUS-SPOMM runs that reflect ontological features in the model
- The mapping isn’t quite as simple as the method suggests, as there are overlaps in behaviour
- This work only looked at time series data, not networks or spatial statistics
- There are a large number of potential variables that we could look at in FEARLUS-SPOMM, and a huge number of statistics for each – this multi-dimensional space is challenging to cluster meaningfully without a huge sample
- Some of the statistics are not computable (lead to infinite or not-a-number results) if, for example, there is no change – reject the run (not complex?) or reject the statistic?