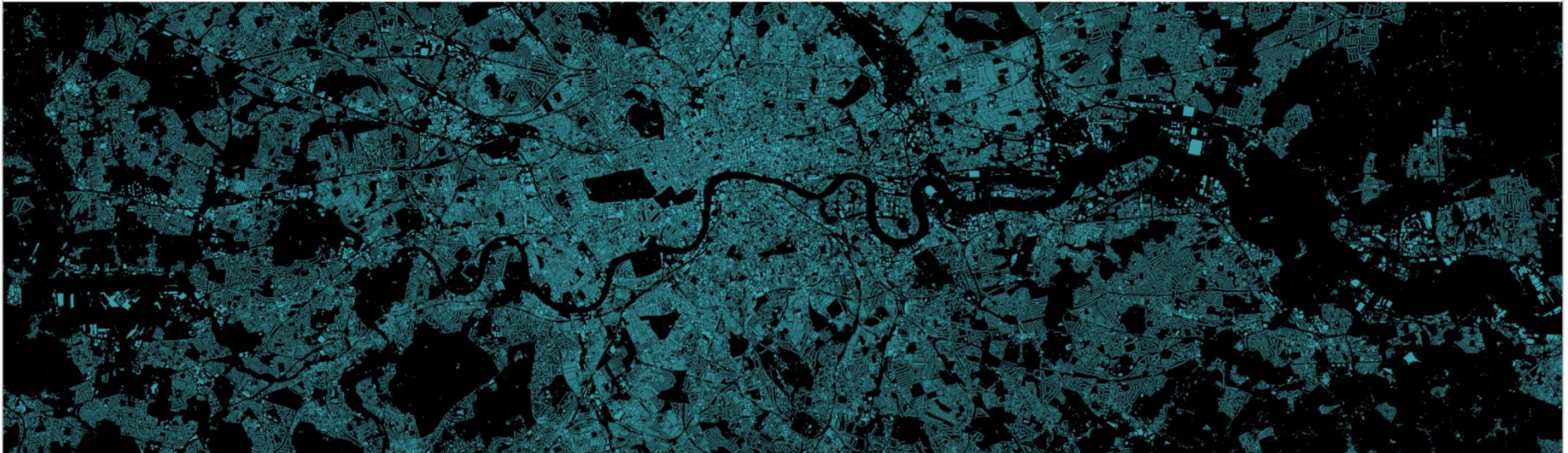# Predictive analysis and probability of error scoring, study and practise

Matt Tobin

Ordnance Survey – Quality

January

Why predictive analysis?

Rise in demand for geospatial data
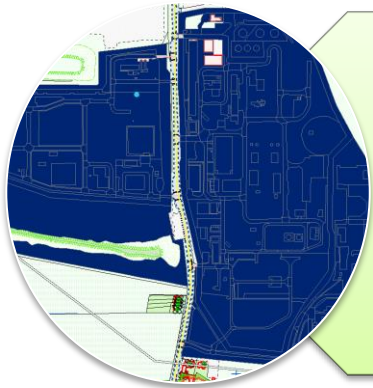
Challenges to current methodologies

Innovation & continuous improvement

Useful Information

What is predictive analysis to us?

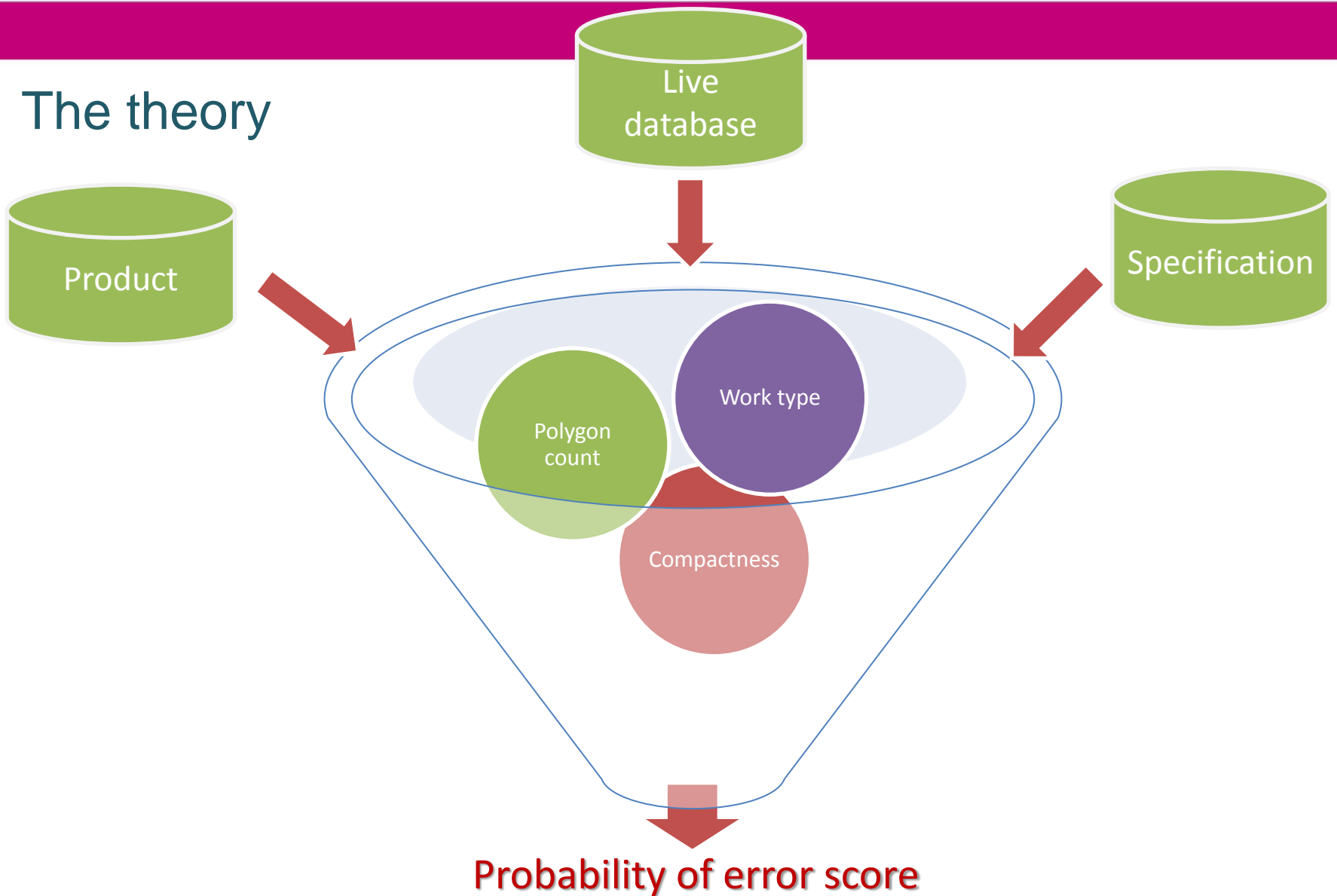Using trends and features to analyse where error is likely to be located.

What is OS Sites layer?

Provision of functional site extent polygon encompassing the perimeter of sites such as; schools, airports, train stations, hospitals and universities. This is not the legal extent.
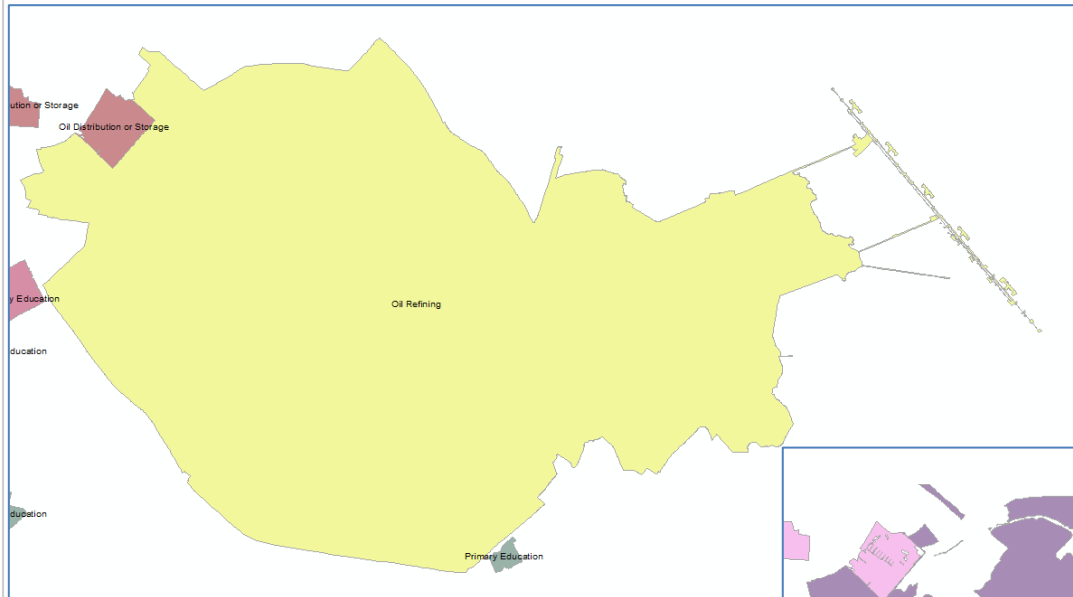
Functional Site Example

The theory

Live database

Product

Specification

Polygon count

Work type

Compactness

Probability of error score

Compactness

❑ A standard geometric measure of the compactness of a shape is given by the formula:
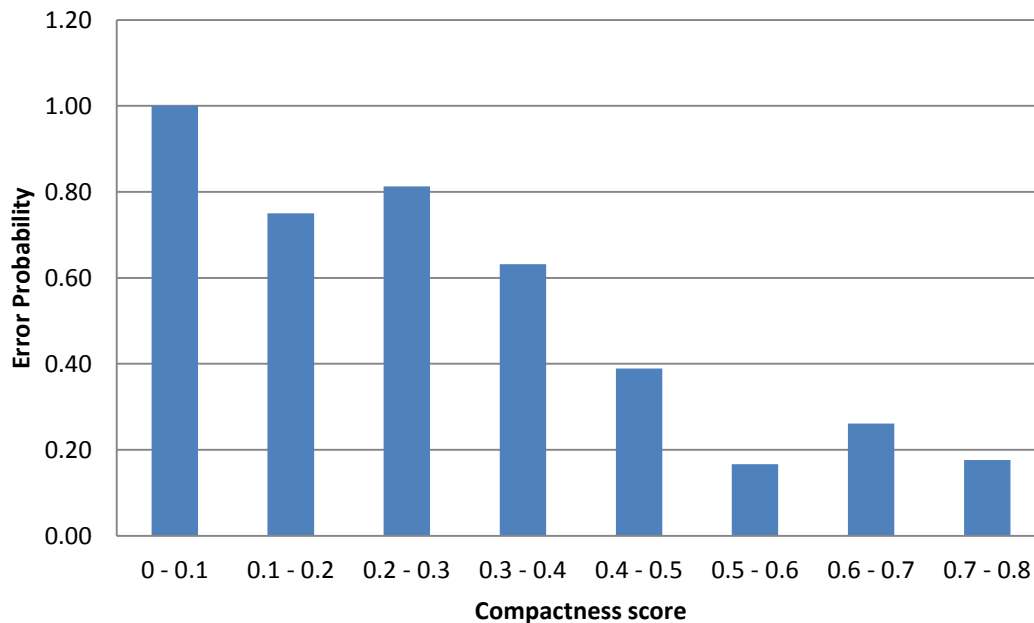
$$\square\frac{\text{Area}}{\text{Perimeter}^2} \times 4\pi$$
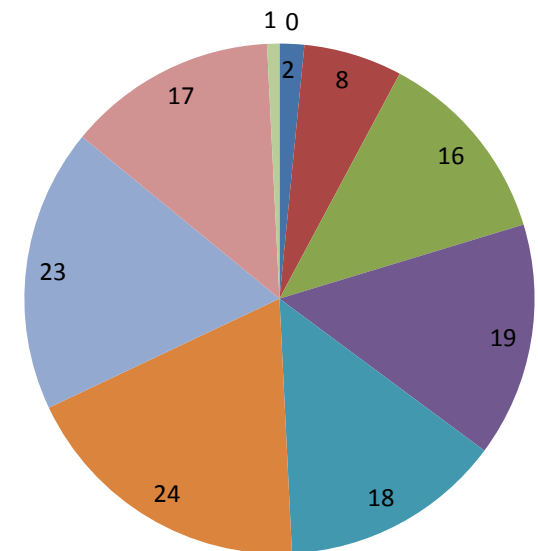
Degradation

Correlation

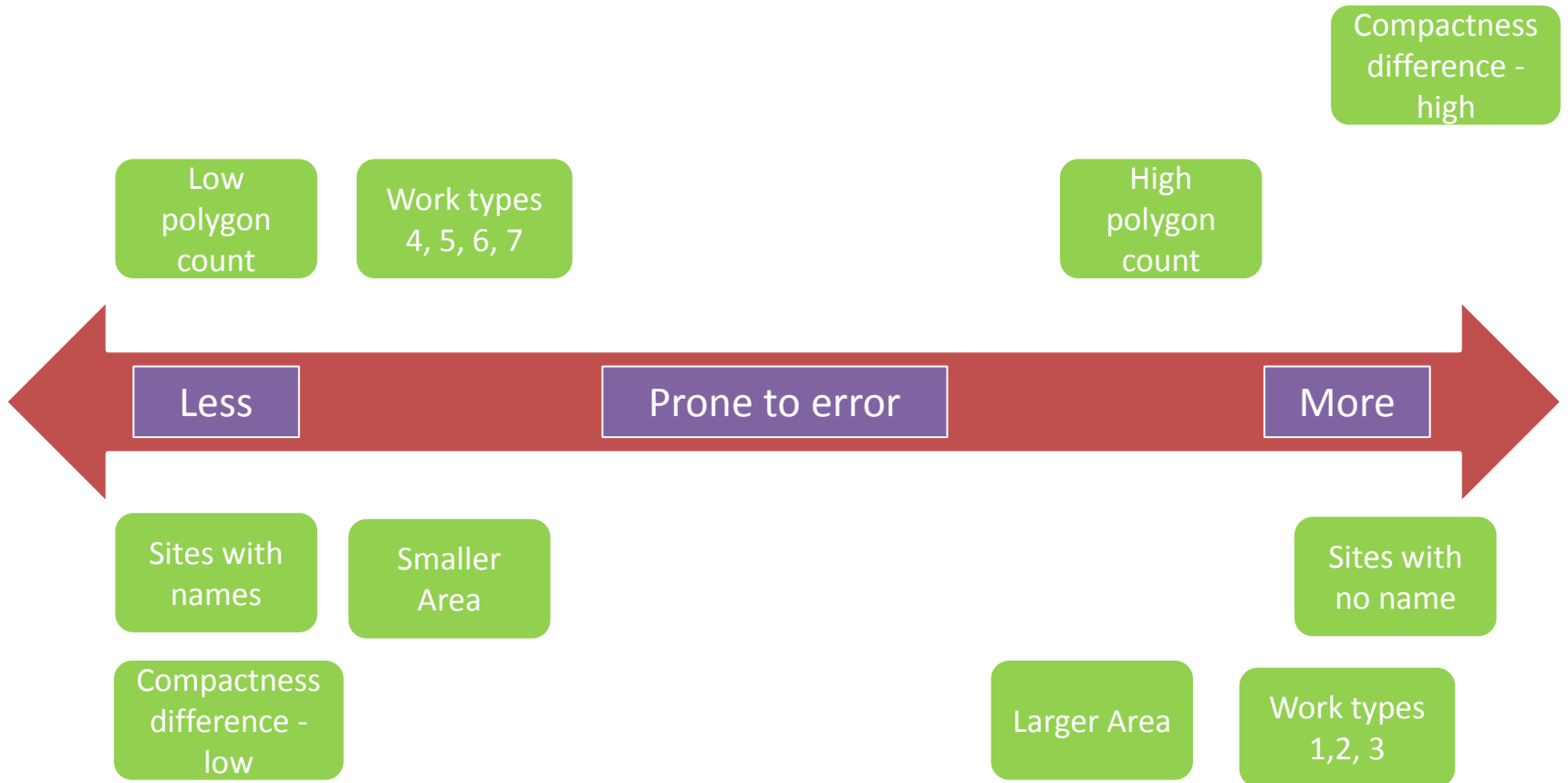**Error probability against compactness score**

Proportion of Functional Sites in each compactness score

❑ High inverse correlation between error probability and compactness score prompting its inclusion in the error probability measure

Methodology - score creation

Equation

$$P(\text{functional site needs checking})$$
$$= f(\text{Compactness score, member count, existance of name, area, work type})$$

$$P(\text{functional site needs checking})_{high} \Rightarrow \text{Compactness score } < 0.5$$
$$\Rightarrow \text{Work type} = 1,2,3$$
$$\Rightarrow \text{No name}$$
$$\Rightarrow \text{Area } \ll 2000$$

Findings

❑ Feedback of error to creator; system, process or editor built in to the method

❑ Decrease in visual scanning

❑ Estimated impact ⅔ of error identified without visual scan

❑ On current measure only 1000 sites to be checked out of the changed 5000 sites

Challenges

Matt Tobin

Matt.tobin@ordnancesurvey.co.uk