

Spatial awareness while sampling

H Reynolds, IN Fabris-Rotelli, T Loots
Department of Statistics, University of Pretoria
u12044700@tuks.co.za

Keywords: spatial sampling, spatial data, environmental sampling

Waldo Tobler was quoted as saying: "Everything is related to everything else, but near things are more related to each other" [?], this leads us to believe that geographical data points possess some sort of correlation which is not necessarily considered when traditionally sampling is used. Spatial data is believed to be one of three types of populations; namely zero-dimensional (point-like data, such as trees in a forest), one-dimensional (linear data, such as a river or stream) and two-dimensional (areal data, such as soil nutrients) [?]. Within the world of classical sampling there are two approaches that can be followed; design-based or model-based sampling [?]. Design-based sampling was previously believed to be irrelevant for spatial data, however Brus and de Gruijter have since proved otherwise. Packages that are available in R will be used to sample from point locations as well as continuous spatial datasets. These packages are **sp** and **spatstat**. Most researchers opt for systematic or stratified spatial sampling as a means to preserve the spatial information. Another package in R, **spcosa**, designs spatial coverage samples and uses variations of k -means clustering to partition the region of interest into compact blocks [?]. In this talk we discuss approaches to sampling spatially using design-based methods, highlighting how incorporating the spatial information in the design of the sampling plan can improve the estimation of characteristics of interest of a population.